HEDONIC ANALYSIS OF PROPERTY MARKETS: THEORY AND APPLICATIONS TO UK CITIES

BY BRETT DAY

A thesis submitted for the degree of Doctor of Philosophy in Economics

University of London

DEPARTMENT OF ECONOMICS

UNIVERSITY COLLEGE LONDON

January 2005

ABSTRACT

This thesis concerns the hedonic analysis of property markets. In particular, it investigates the extent to which such analysis can reveal household preferences for the avoidance of exposure to transport-related noise pollution. The thesis is divided into three Parts.

Part 1 provides a thorough exposition of the economic theory of property markets. It contains two chapters. The first details the establishment of a market-clearing hedonic price equilibrium. The second outlines how the property market model can be used to identify measures of welfare change resulting from exogenous changes in environmental quality. As well as providing possibly the most complete and coherent exposition of this expansive and occasionally confused literature, Part 1 also contributes new insights into welfare measurement when landlords are constrained in their responses to environmental change.

The following two Parts of the thesis concern empirical applications of hedonic analysis to property markets. Part 2 is concerned with the estimation of hedonic price functions for the City of Birmingham property market. The unique innovations presented here include the application of techniques for partitioning data in order to improve specification of the hedonic price function and the application of semiparametric estimators in order to redress spatial autocorrelation amongst regression residuals.

Finally Part 3 of the thesis concerns itself with welfare analysis. Specifically, it provides a thorough discussion of the implications of the theory in Part 1 for empirical estimation of preference parameters. Following these empirical guidelines and drawing on results from Part 2, welfare estimates for changes in exposure to traffic-related noise pollution are provided. As far as the author is aware, these are the first welfare estimates for noise pollution to be derived from a hedonic analysis in a theoretically consistent manner.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

My first acknowledgement, as is traditional in such circumstances, is to my supervisor Professor David Pearce. In this case, I feel much more than an acknowledgement is due. Professor Pearce, was mostly responsible for my decision to pursue a career in environmental economics, he provided me with my first opportunity to set sail on such a career and since then has been almost wholly responsible for ensuring that that wobbling tub has not capsized. I hope that this thesis provides some small measure of reassurance that the faith he has shown in me has not been entirely unfounded. Moreover, I believe this to be the last PhD thesis that Professor Pearce intends to supervise. This is a sad loss. His keen insight, encyclopaedic knowledge, unwavering support and tireless enthusiasm have been an inspiration to me and a whole generation of young environmental economists.

The research described in this thesis would not have been possible without the input of Professor Ian Bateman and Dr. Iain Lake. In particular, Dr. Lake compiled the superb data sets that form the basis of the empirical analyses described in Parts 2 and 3 of this thesis. The brief descriptions of these data sets that I have included in this document do scant justice to the painstaking quality of Dr. Lake's work. To Professor Bateman, I offer my heartfelt thanks for his support, encouragement and friendship. His has been the most active academic input in the writing of this thesis.

Of course, these acknowledgements would be incomplete without mention of my family and friends. I look forward to a future in which conversations with my parents no longer include the dreaded question; "How's the thesis going?". Many thanks Mum and Dad, this would have been impossible without you. Also to Claire, who was rudely dragged from her spiritual home and dumped in the back end of beyond. Despite that horror she has stuck by me, putting up with the endless late nights and periodic bouts of stress and desperation. My love and thanks.

Finally, my thanks to Professor Joe Sweirzbinski and to all my friends and colleagues from CSERGE at UCL; it's taken some time but I got there in the end! Also, I gratefully acknowledge the research grants provided by The Scottish Office, Department for Transport and the Economic and Social Research Council that have made this work possible.

PART 1

THE THEORY OF HEDONIC ANALYSIS OF PROPERTY MARKETS

CHAPTER 1: THE HEDONIC MODEL OF PROPERTY MARKETS

1. Introduction

One of the most familiar models in economics is that of price determination in the market. The market for a particular good consists of a large number of consumers whose demand for the good is met by the production of a large number of firms. The market mechanism works to reconcile the needs of consumers and firms by establishing the price at which aggregate demand is equal to aggregate supply and the market clears. At this price the market is said to be in equilibrium since there is no excess demand for the good and firms cannot increase their profits by changing their production of the good.

For many goods, however, this simple model is inadequate. For example, the simple model predicts that once in equilibrium the market will determine one price for the good. However, in a market such as that for housing we observe different properties commanding different prices. Indeed, housing is an example of a *differentiated good*. Such goods consist of a diversity of products that, while differing in a variety of characteristics, are so closely related in consumers' minds that they are considered as being one commodity. Many other goods, including breakfast cereals, cars and beach holidays might also fit this description.

Though the simple model does not adequately explain the workings of markets in differentiated goods, it would appear that a similar market mechanism is in operation. Market forces determine that different varieties of the product command different prices and that these prices depend on the individual products' exact characteristics. For example, properties that have more bedrooms will tend to command a higher price in the market than properties that have fewer bedrooms. Furthermore, the set of prices in the market would appear to define a competitive equilibrium. That is, in general, the market will settle on a set of prices for the numerous varieties of the differentiated good that reconcile supply with demand and clear the market.

In a seminal paper, Rosen (1974) proposed a model of market behaviour that described the workings of markets for differentiated goods. The model that Rosen presented provides the theoretical underpinnings for hedonic valuation and will provide the subject matter for the first two chapters of this thesis.

2. The Property Market: The Differentiated Good

Property markets tend to be defined spatially. We shall assume that at any point in time, all of the properties in one urban area represent the products in the property market. The *households* wishing to live in these properties represent the consumers in this market and the *landlords* that own the properties represent the producers in this market.

Clearly the set of properties in the market represent a differentiated good. We could describe any particular property by the qualities or characteristics of its structure, environs and location; that is by the vector,

$$z = (z_1, z_2, \ldots, z_K), \tag{1}$$

where z_i (i = 1 to K) is the level or amount of any one of the many characteristics describing a property. Indeed, the vector z completely describes the services provided by the property to a household.

For the sake of simplicity let us assume that the z_i are measured in such a way that we can consider them as "goods" as opposed to "bads". For example, one of the characteristics of a property will be its exposure to road noise. Rather than measuring this as the level of "noise", we can simply invert the scale and measure it as the level of "peace and quiet".

Further, let us assume that the set of properties in the market is fixed. That is we assume that in the short-run no new properties are built. That is not to say that the characteristics of properties do not change. A landlord maintains the quality of the property by constant renovation and maintenance. Alternatively the landlord can improve the quality of the property through investment. Building an extension, converting a loft or basement, installing double-glazing or central heating, improving the quality of the décor, indeed carrying out any number of alterations and improvements can increase the values of certain of the characteristics of the property. On the other hand, disinvesting, that is failing to maintain and renovate the property, will lead to the quality of certain of its characteristics declining. Of course, certain characteristics of the property cannot be influenced by the actions of the landlord. Most notably the landlord has little influence over the characteristics of the property that are location specific such as its proximity to places of work or to local amenities, or the property's exposure to noise and air pollution.

When households select a particular property in a particular location they are selecting a particular set of values for each of the z_i. We can imagine this market for properties as being one in which the consumers consider a variety of somewhat dissimilar products which differ from each other in a number of characteristics including, amongst many characteristics, number of rooms, size of garden, distance to shops and environmental characteristics such as levels of pollution or noise.

3. The Property Market: The Hedonic Price Function

The price of any property will be determined by the particular combination of characteristics it displays. Naturally we would expect properties possessing larger quantities of good qualities to command higher prices and those with larger quantities of bad qualities to command lower prices. We encapsulate the relationship between a properties price and the vector of values describing its characteristics (z) in the function;

$$P = P(z) \tag{2}$$

where P is the price of a property and $P(\cdot)$, is known as the *hedonic price function* (HPF); 'hedonic' presumably because it is determined by the different qualities of the differentiated good and the 'pleasure' these would bring to the purchaser.

In the property market this price is the rental that a household pays to the landlord. In effect, every household in the urban area is purchasing the flow of services derived from the characteristics of the property per period of time. To clarify $P(z)$ is the per period payment made by a household to a landlord for the use of a property over that period.

Of course, many households own their own homes. In this case we treat homeowners as landlords that rent from themselves. If markets are operating perfectly, and generally we assume that they are, then the price at which the household purchases the property will be the discounted sum of all the future per period rents from that property according to;

$$Purchase\ Price = \sum_{t=1}^{T} \frac{P(z)}{(1+d)^t} \qquad (3)$$

where t indexes each time period, T is the expected life of the property and d is the discount rate. Equation (3) allows purchase prices to be translated into per period rentals a relationship we shall exploit in the empirical investigation described in Chapter 9 of this thesis.

One feature of the HPF that warrants further discussion results from the fact that households are unable to "repackage" the differentiated goods. In other words, households cannot break up the differentiated good into its constituent parts and enjoy the benefits of each characteristic separate from the whole. For example, talking in terms of just one characteristic, two houses with one bedroom are not equivalent to one house with two bedrooms since a household cannot live in both properties simultaneously. Similarly, renting a property with four bedrooms for half a year and a property with two bedrooms for the other half is not the same as renting a three-bedroom house all year round. In markets for ordinary goods opportunities for arbitrage ensure that the price for each unit of a product (its marginal price) is constant. That is one can expect to pay the same price for each unit no matter whether it is the first, second or hundredth unit purchased. Since the "products" that make up a property are bundled in such a way that they cannot be repackaged, in the short run at least, arbitrage activity is precluded in the housing market. Market forces do not work to ensure that the marginal price of bedrooms is constant.

This observation leads to two interesting insights.

- *Marginal prices may not be constant.* Indeed, more typically, the additional amount paid for properties enjoying increasingly higher quantities of a characteristic (the marginal price of that characteristic) declines as the total level of that characteristic increases.

- *The price of one characteristic may depend on the quantity of another.* As an example, a house with a garden is more desirable than a house without. Further, if the aspect of the house is north-south, having a garden may be even more desirable since it will enjoy longer exposure to the sun. Now consider the extra paid for a north-south aspect, effectively the 'price' of north-south aspect. Without a garden, north-south aspect may be somewhat desirable, but households are unlikely to pay a great deal more for a property with this characteristic compared to an identical property with east-west aspect. For properties with a garden, on the other hand, aspect may be a much more important consideration. It would not be surprising that the price of north-south aspect will depend on whether a property has a garden or not.

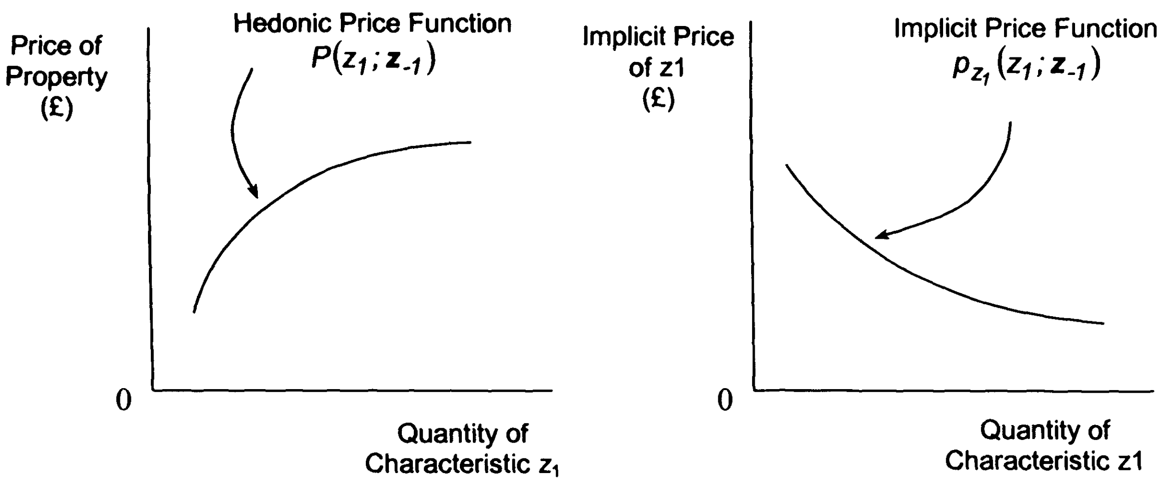
These two observations will be important in empirical applications that attempt to estimate the HPF from market data such as those discussed in Part 2 of this thesis.

To illustrate the HPF, consider the illustration in the left hand panel of Figure 1. Plotted on the vertical axis is the price (rental per unit time) of property. Along the horizontal axis is quantity of a particular housing characteristic labelled z_1. Further, let us introduce some new notation, z_{-1}, which is the vector containing the levels of all property characteristics barring z_1. Notice that in the HPF in Figure 1, z_{-1} comes after a semicolon, indicating that, for the purposes of illustration, these other characteristics are held constant at some given level.

In this hypothetical case, the HPF rises from left to right implying that the greater the quantity of characteristic z_1 the higher the price commanded by a property in the market. Notice also that the marginal price of characteristic z_1 is not constant. The slope of the curve becomes progressively flatter and the incremental increase in a property's market price resulting from its possessing more of characteristic z_1 declines.

It may be easier to illustrate the idea of non-constant marginal prices through actually plotting the additional amount that must be paid by any household to move to a bundle with a higher level of that characteristic, other things being equal. This is illustrated in the right hand panel of Figure 1. This new function is known as the *implicit price function*; *implicit* because the marginal price of a characteristic is revealed to us indirectly through the amounts households are prepared to pay for the whole property of which the particular characteristic is only a part.

Figure 1: The Hedonic Price and the Implicit Price Schedules for characteristic z_1



Mathematically, the implicit price is derived as the partial derivative of the HPF (Equation 2) with respect to one of its arguments, z_i, according to:

$$p_{z_i}(z_i, \mathbf{z}_{-i}) = \frac{\partial P(\mathbf{z})}{\partial z_i} \tag{4}$$

Again $p_{z_i}(z_i, \mathbf{z}_{-i})$, the marginal price function of characteristic z_i, does not have to be a constant. We have dwelt on the subject of non-constant marginal prices for characteristics since, as we shall see in Chapter 8 of this thesis, they are the source of much of the complication that confounds the empirical estimation of welfare measures using hedonic analysis.

4. The Property Market: Household Choice

Let us take as a fact that the HPF, $P(\mathbf{z})$, emerges from the interaction of households (demanders) and landlords (suppliers) and represents a market clearing equilibrium. We shall return to the mechanism by which this equilibrium is derived, but for now let us focus on how households facing such a hedonic price schedule determine their optimal residential location.

The model that Rosen developed to explain these decisions is based on a number of assumptions. Amongst these, some of the most important are that;

- Each individual household in the market is a price taker; they make their choice of location based on the hedonic price schedule they observe in the market and cannot influence this schedule through their actions. As articulated by McConnell and Phipps (1987) and Palmquist (1991) amongst others, this assumption allows us to ignore the supply side of the market in modelling households' residential decisions. Given the size of the urban property markets in which hedonic pricing techniques are usually applied, such an assumption would appear reasonable.
- Each household only purchases or rents one property[1]. If households purchase a second home, say a holiday home, then this should be considered as a separate good being purchased in a separate hedonic market. Again, this assumption is, in general, readily defensible.

Given these assumptions, Rosen sets out a model in which households choose their residential location so as to get the maximum flow of benefits. To do this, it is assumed that households in the market have well-defined preferences over all goods and that these preferences can be represented by the utility function;

$$U(z, x; s) \tag{5}$$

The utility function contains three arguments;

- z which represents the levels of the different characteristics of property that a household could purchase or rent.

[1] Further, if households also act as landlords to other households, then their decisions concerning their own choice of residential location are assumed to be independent of their decisions concerning these other properties.

- x which represents all other goods outside the property market. As a matter of convenience we standardise x to have a unit price, such that, we are effectively representing all other goods by a quantity of money[2].

- s which represents the characteristics of the household themselves. Clearly, the quantity of utility a household enjoys from any of the other arguments will depend on their own characteristics. For example, having a swimming pool in the back garden will confer little benefit on a household of non-swimmers.

For now we won't specify the exact form of the utility function.[3] For our purposes, we can continue by assuming there is such a function and that it is the same for each household conditional upon their characteristics. At the end of this Chapter we shall discuss a number of models that have been developed using specific assumptions concerning the form of the utility function.

Households choose levels of z and x to maximise $U(z,x;s)$ subject to the constraints imposed upon them by their budget. Since the price of x is taken as unity, the budget constraint can be represented as;

$$y = x + P(z) \tag{6}$$

where y is household income per period.

As with the standard consumer choice problem, we can use Equations (5) and (6) to set up the Lagrangian Function;

$$L = U(z,x;s) + \lambda(y - x - P(z)) \tag{7}$$

Maximising this with respect to x, z and the Lagrange multiplier λ gives rise to the first order conditions;

---

[2] In the economics literature x is referred to as a numeraire, a composite good or a Hicksian bundle.

[3] We do however assume that the utility function is strictly increasing in the arguments x and z (remember we have already assumed that all attributes were measured as goods). Further, for mathematical simplicity, we assume that the utility function is strictly quasiconcave and twice continuously differentiable.

$$\frac{\partial L}{\partial z_i} = U_{z_i} - \lambda P_{z_i} = 0 \qquad (i = 1 \text{ to } K) \qquad (8)$$

$$\frac{\partial L}{\partial x} = U_x - \lambda = 0 \qquad (9)$$

$$\frac{\partial L}{\partial \lambda} = y - x - P(z) = 0 \qquad (10)$$

Where;

- U_x is the partial derivative of the utility function with respect to the composite good. As such, U_x can be interpreted as the extra utility that comes from an extra unit of money, all else being equal.
- U_{z_i} is the partial derivative of the utility function with respect to property characteristic z_i. This represents the extra utility that comes from choosing a property with one extra unit of characteristic z_i, all else being equal.
- P_{z_i} is the partial derivative of the HPF with respect to property characteristic z_i. Of course this is simply the implicit price function for characteristic z_i as presented in Equation (4). Indeed, $P_{z_i} = p_{z_i}(z_i, z_{-i})$.

Equations (8), (9) and (10) are the conditions that define the household's optimal choice of residential location. That is, given the constraint of their budget, the flow of utility that the household enjoys will be maximised by choosing a property whose characteristics simultaneously satisfy the conditions laid out in Equations (8), (9) and (10).

In their present form these conditions provide us with little insight into the household's choice behaviour. However, if we rearrange Equations (8) and (9) and divide one by the other (thereby eliminating the Lagrange multiplier) we reveal that one of the conditions for optimal choice is given by the expression;

$$\frac{U_{z_i}}{U_x} = p_{z_i}(z_i, z_{-i}) \qquad (i = 1 \text{ to } K) \qquad (11)$$

To illustrate the condition laid out in Equation (11) Rosen defined a function that he termed the *bid function*, whose slope is given by the ratio of marginal utilities, U_{z_i}/U_x. More usually we would expect to see the ratio of two marginal utilities preceded by a negative sign. In such a case, the expression would represent a *marginal rate of substitution*, e.g. $-U_{z_i}/U_x$; the quantity of one good that a household is willing to give up in order to obtain one more unit of another good such that their overall well-being does not change. In the same way that U_{z_i}/U_x defines the slope of Rosen's bid function, the marginal rate of substitution defines the slope of an indifference curve. In hedonic analysis there is a simple correspondence between the indifference curve and the bid function that goes some way towards clarifying the nature of the latter.

Let us spend a little time considering indifference curves. In mathematical terms, the indifference curve is implicitly defined as;

$$U(z, x; s) = u \tag{12}$$

Where u is any specified level of utility. Thus, the indifference curve depicts combinations of x and z that confer the same level of well-being or utility on the household. Indeed, solving Equation (12) for x would give us a general expression for an indifference curve that we can denote;

$$x = x(z; s, u) \tag{13}$$

The left hand panel of Figure 2 shows a set of indifference curves between x (the quantity of money to spend on other goods) and z_1 (one of the attributes of a property)[4]. Each indifference curve depicts combinations of x and z_1 that confer the

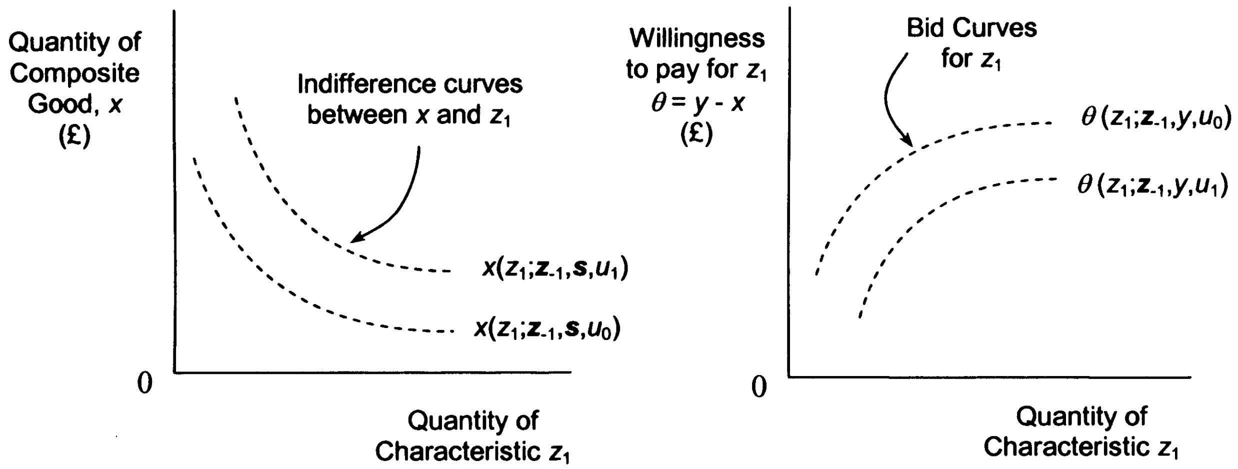
---

[4] For diagrammatic exposition, it is necessary to present indifference curves in terms of only one property attribute. The assumption in Figure 2 is that all other attributes are held at some constant level. In reality, the indifference 'curve' would be a multidimensional indifference surface plotting combinations of x and quantities of each of the attributes in z between which the household is indifferent.

same level of well-being or utility on the household.[5] As such, the slope of the indifference curve gives the rate at which households are prepared to give up money for other goods in order to acquire more of the housing attribute whilst not changing their overall well-being. That is, the slope of the indifference curve is the marginal rate of substitution between x and z_1; $-U_{z_1}/U_x$.

Any combination of x and z_1 that lies above and to the right of an indifference curve provides the household with more money to spend on other goods and/or more of the housing attribute. By definition they must gain more utility from such a bundle than from any bundle lying on the indifference curve. Consequently the higher indifference curve in Figure 2 identifies bundles of x and z_1 that confer more utility on the household, u_1, than bundles lying along the indifference curve for u_0.

Figure 2: Indifference Curves and the Bid Function



So far we have not considered the fact that the household is constrained in their choices of x and z by their limited budget, y. Money spent on other things is money that cannot be spent on housing attributes. Let us define, therefore, an amount that we shall call a *bid* as;

$$\theta = y - x \tag{14}$$

[5] The "classic" downward sloping shape for the indifference curves in Figure 2 stems from our assumptions concerning the utility function described in footnote 3.

That is the bid, θ, represents the total amount a household could pay for a property given that they spent x on other goods. Clearly the relationship between the bid, θ, and the amount spent on other goods, x, is very simple; as one goes up by a certain amount the other falls by the same quantity[6]. Indeed, using Equation (14) we could redefine the indifference relationships of Equation (13) in terms of bids rather than money spent on other goods. Replacing Equation (13) in Equation (14) gives;

$$\begin{aligned} \theta &= y - x(z;\ s,u) \\ &= \theta(z;\ y,s,u) \end{aligned} \tag{15}$$

The bids defined by Equation (15) are a special type of bid. They are bids for a property with characteristics z that result in the level of utility u. Indeed, Equation (15) defines Rosen's *bid function.*

In words, the bid function depicts the maximum amount that a household would pay for a property with attributes z such that they could achieve the given level of utility, u, with their income, y. Notice that increases in income translate directly (i.e. pound for pound) into increases in the bid function.

The bid function can be illustrated as *bid curves* as depicted in the right hand panel of Figure 2. Notice that bid curves still define indifference relationships. They depict combinations of property attributes, z, and payments for those attributes, θ, between which the household is indifferent. Accordingly, all combinations on a particular bid curve provide the household with the same level of utility. Combinations lying below and to the right of a particular bid curve represent bundles providing more attributes and/or lower payments. Consequently the lower bid curve in Figure 2 provides the household with greater overall utility, u_1, than that provided by the higher bid curve, u_0.

Since the bid curve is, roughly speaking, an inverted indifference curve, the slope of the bid curve will be the same as the slope of the indifference curve but with the

[6] Though, clearly, the bid is constrained in that it cannot be greater than income, y.

opposite sign i.e. U_{z_i}/U_x. And, of course, this ratio represents the left hand side of the condition for optimal residential location given in Equation (11).

So far our analysis has defined two closely related functions the indifference curve (Equation 13) and the bid curve (Equation 15). As yet, however, we have not determined how these functions are important in defining households' optimal residential choice. To do this, we must make use of the last of the first order conditions, that defining the budget constraint (Equation 10).

To plot the budget constraint we must rearrange Equation (10) to give;

$$x = y - P(z) \tag{10a}$$

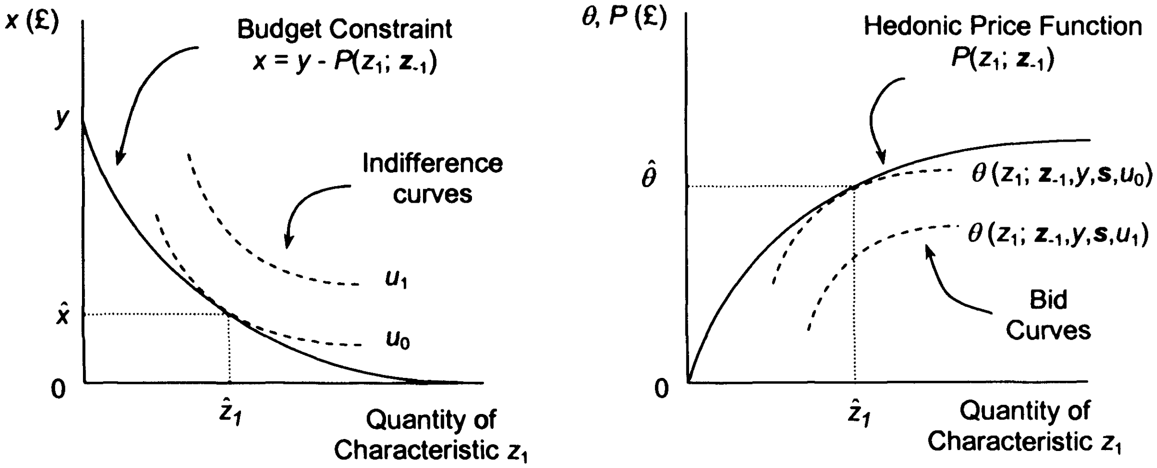
In the left hand panel of Figure 3, the budget constraint has been added to the indifference diagram. The budget constraint describes all combinations of x and z_i that the household is able to buy given their income, y. Bundles that are on or below the budget constraint are affordable to the household. In order to maximise their utility the household must choose the bundle amongst affordable combinations of x and z_1 that lies on the highest indifference curve. The optimal bundle ($\hat{x}$, $\hat{z}_1$) will be defined as the point of tangency between this highest indifference curve and the budget constraint. Notice that throughout this Chapter we use a hat to represent a chosen bundle. Any other bundle in the affordable set will lie on an indifference curve that provides a lower level of utility.

The left hand panel of Figure 3 will be familiar to students of economics. However, the diagram differs from the usual consumer choice problem in that the budget constraint is not linear since, as we have already established marginal prices may not be constant in hedonic analysis.

The choice of an optimal bundle of housing attributes can just as easily be presented in terms of bid functions. Of course we have to transform the constraint to be expressed in the same terms as bids. Remember that the vertical axis of the bid function graph is measured in terms $y - x$; that is money available to spend on housing attributes. Rearranging the income constraint, Equation (10), gives;

$$y - x = P(z) \tag{10b}$$

Figure 3: Choice of Optimal Residential Location using Indifference Curves and the Bid Function

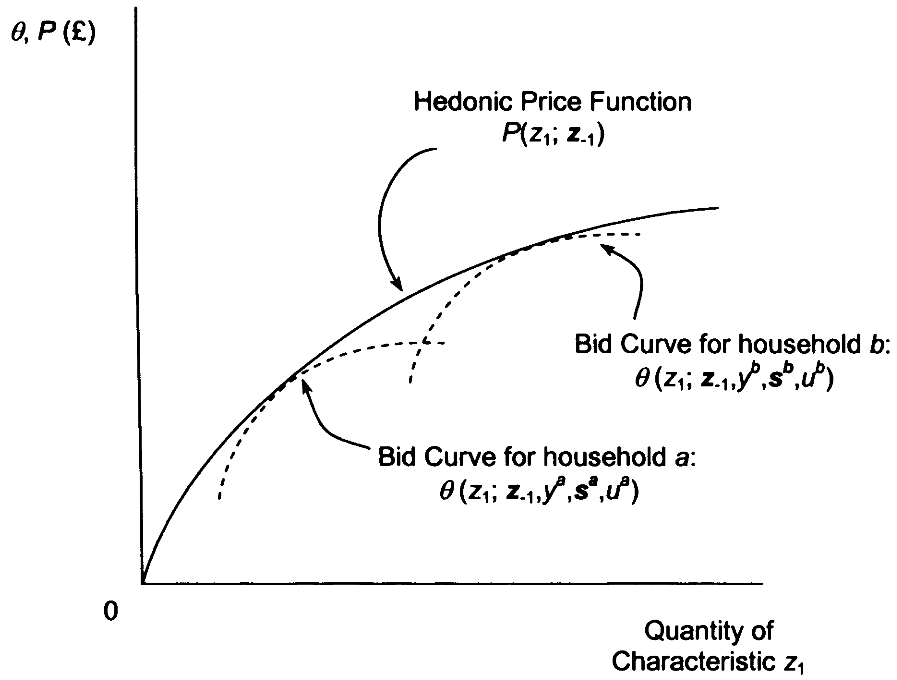


In other words, the relevant constraint is simply the HPF. When presented in this manner, the maximisation problem amounts to the household choosing the bundle of housing attributes that positions them on the bid curve providing the highest level of utility whilst remaining compatible with market prices. Again this can be illustrated as a point of tangency as depicted in the right hand panel of Figure 3. Indeed, the point of tangency between the lowest bid curve and the HPF defines the bundle of housing attributes that fulfil the first order conditions for an optimal choice (Equations 7, 8 and 9).

Of course, households do not have the same income nor do they have the same socioeconomic characteristics. Since both these arguments enter the bid function, we would expect bid curves to differ across households with different characteristics. Bid curves for two different households, denoted a and b, are illustrated in Figure 4. Notice that the conditioning arguments, y and s, are superscripted with this household indicator, showing that their values are specific to the particular household.

Again the optimal choice of property for each household will be defined by the tangency of a bid curve with the HPF. Since the bid curves for the two households are different, the attributes of the property defined by this point of tangency will also differ. Notice that the utility level, u, that defines the optimising bid curve is also superscripted by a household indicator. Clearly, there is no reason to expect that the level of utility achieved by the two households would be the same.

Figure 4: Choice of Property Attributes for Different Households



If we were to add to Figure 4 the optimising bid curves for all the households in the market we would find that they were all tangential to the HPF. Variation in household characteristics would mean that these points of tangency defined properties with different levels of the various housing attributes. In the terminology of economics, the HPF forms an upper envelope to these optimising bid functions.

5. The Property Market: Landlord Choice

So far we have examined the property market solely from the demand side; that is, in terms of households choosing between differentiated products. Though this decision is of greater interest for the research objectives of this thesis, it is worth taking some time to examine the supply side of the market; that is, to describe how landlords make their decisions concerning the type of properties to supply.

To simplify the analysis let us assume that each landlord rents out only one property.[7] In each period of time a household pays the landlord rent in order to live in this property. Of course this rent does not represent pure profit to the landlord. The landlord incurs costs in supplying this property for rental;

- First and foremost, through the initial purchase of the property. [8]
- Second, through maintaining the quality of the property by constant renovation and maintenance.
- Finally, through investments or disinvestments designed to change the attributes of the property subsequent to its purchase.

To incorporate these costs into the "per period" model, all discrete investments must be converted to equivalent per period costs. For example, the purchase price of the property can be expressed as an equivalent series of per period payments using Equation (3).[9] In the same way, it is possible to express a discrete investment in the property, say the installation of double-glazing, as the discounted sum of a series of smaller equal-sized costs made over the expected lifetime of the investment.

The per period cost to the landlord of supplying a property with characteristics z, is given by the cost function[10];

$$C\left(z;\, P^{t}(\hat{z}), \underline{z}, r\right) \tag{16}$$

[7] The analysis remains relatively simple if we assume that the landlord rents out more than one property but that each property has the same characteristics. However, if the landlord rents out several properties with differing characteristics the model becomes considerably more complex whilst adding little to our understanding of the workings of the hedonic market.

[8] Or even the initial construction of the property.

[9] Indeed, in the UK, this does not represent a major abstraction from reality. It is typical for landlords to take out a mortgage in order to purchase a property. The purchase price of the property (plus interest on money borrowed) is repaid, therefore, in a series of monthly instalments.

[10] This cost function is the result of a minimisation problem in which the landlord attempts to find the cheapest cost means by which to produce a property with characteristics z.

The cost of producing a property with characteristics z will differ across landlords for a number of reasons. The factors determining these differences are captured in the three conditioning arguments entering the cost function. These are;

- $P'(\hat{z})$, which defines the *price paid for the property when first purchased.* As we shall discuss later in this Chapter, the HPF is assumed to adjust to changes in market conditions. Clearly, the cost of supplying a property will depend on the market prices reigning at the time of purchase, $P'(\cdot)$. Since this cost (expressed in equivalent per period terms) is constant for each landlord and independent of changes in the property market or in the characteristics of the property, we suppress this argument in the cost function henceforth.

- the vector $\underline{z}$, which defines the levels of attributes that, following the initial property purchase, are *provided costlessly to the landlord.*

    - For *structural attributes* this vector is likely to be identical to $\hat{z}$, the vector of housing attributes purchased by the landlord. For example, having purchased a two-bedroom house, the landlord does not have to pay further in order to maintain this number of bedrooms in the property.

    - For *locational, neighbourhood and environmental attributes*, the levels of $\underline{z}$ will tend to be determined by *exogenous* factors. For instance, $\underline{z}$ would include a measure of the level of crime in the area. This level of crime is provided costlessly to the landlord in so much as it is determined by government spending on crime prevention.[11] Clearly, the cost to the landlord of providing a certain level of security at the property will be highly influenced by this baseline level of crime.

      Other examples include the property's proximity to recreational facilities, its access to public transport, levels of air pollution and levels of noise pollution. Indeed, to a large extent, the values of $\underline{z}$ for non-structural attributes of the property are determined by public policy. Policies that

[11] Of course, the landlord indirectly pays for such public goods through taxation, but payments are not directly linked to the level of the attribute enjoyed at the property and payments cannot be unilaterally altered so as to influence this level.

reduce crime, redirect traffic, combat air pollution or increase the quality of public transport will determine the values included in $\underline{z}$ for certain attributes. As we shall discuss in the next chapter, the primary aim of this thesis is to use hedonic analysis so as to determine the benefits to households and landlords of public projects that change the levels of these housing attributes.

- the vector $\boldsymbol{r}$, captures *other parameters* important in determining the landlord's costs. For example, $\boldsymbol{r}$ will include the characteristics of the landlord and the market price of investments. To illustrate, a landlord that is a capable plumber may be able to improve the quality of a property by installing an electric shower unit. The costs may be lower to this landlord than to another who has to seek professional help to achieve the same improvement.

The cost function, therefore, determines the per period cost of supply of a property with characteristics z, given the purchase price of the property $P^{t}(\hat{z})$, the levels of exogenously determined property attributes, $\underline{z}$, and a number of other parameters including the characteristics of the landlord, $\boldsymbol{r}$.

Importantly, landlords have the ability to change the characteristics of their property. For example, a landlord may choose to increase the peace and quiet at the property. For the sake of argument let us assume that the least cost method of achieving this increase is to install double-glazing. In this case the difference in the cost function evaluated at the original level of peace and quiet and the cost function evaluated at the increased level of peace and quiet would be the cost of the double-glazing.[12]

Given the per period cost defined by the cost function, the profit that a landlord derives from renting a property with characteristics z, will be determined by the rental price the landlord can charge for such a property in the market. Hence;

---

[12] An investment such as the installation of double-glazing would require a one-off payment. Of course, it is possible to express this one-off payment as the discounted sum of a series of smaller equal-sized costs made over the expected lifetime of the investment. The increase defined by the cost function would be equal to this extra per period cost.

$$\pi(z;\underline{z},\boldsymbol{r}) = P(z) - C(z;\underline{z},\boldsymbol{r}) \tag{17}$$

Where $\pi(\cdot)$ is the profit function defining the landlord's profit per period.

To make our analysis compatible with that for the demand side of the market, let us define a function that joins all combinations of z and $P(z)$ that return the same profit for the landlord. To do this, set $\pi(z;\underline{z},\boldsymbol{r})$ equal to the constant π in Equation (17). Then solve for the market prices that would be required in order to realise the profit π for different levels of z. Mathematically, this amounts to;

$$\phi(z;\underline{z},\boldsymbol{r},\pi) = \pi + C(z;\underline{z},\boldsymbol{r}) \tag{18}$$

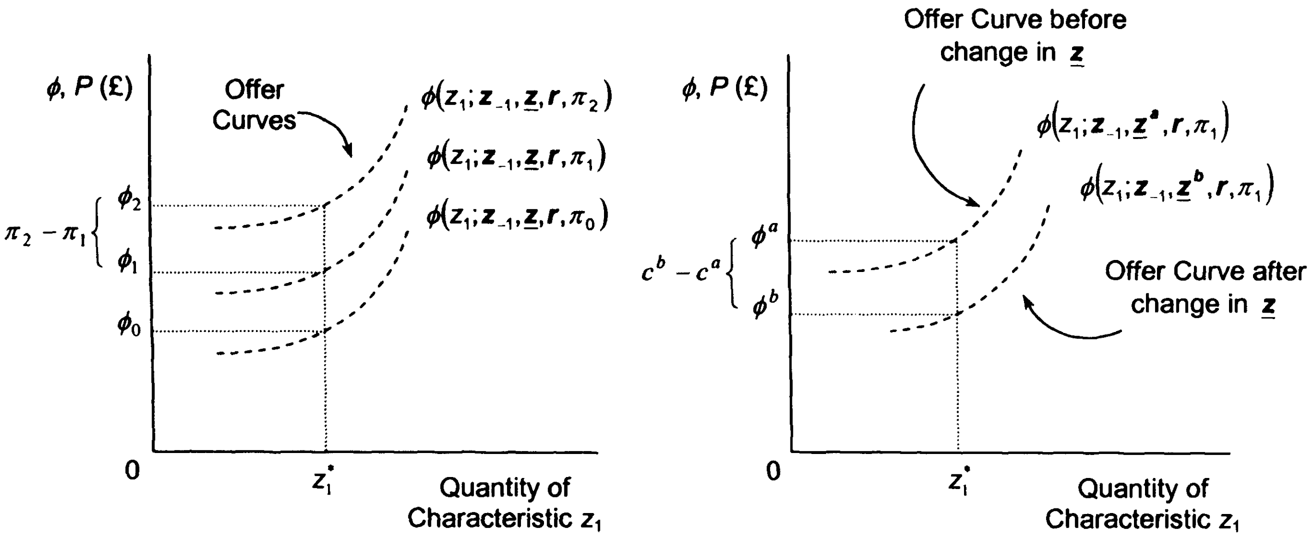
This function is what Rosen terms the *offer function*. The offer function describes the rent the landlord would need to receive in order to achieve a profit of π if he were to provide his property with a level of characteristics given by the vector z.

The offer function can be illustrated as *offer curves*. Each offer curve combines rental prices and levels of attribute provision that result in the same level of profit. The left hand panel of Figure 5 plots one landlord's offer curves for attribute z_1. The upper offer curve represents combinations of attributes and prices that would return a profit of π_2, the middle curve combinations giving a profit of π_1 and the lower curve a profit of π_0.

As we would expect, higher offer curves define higher levels of profit for the landlord. Take for example one particular level of provision of z_1, let us say z_1^*. At this level of provision, the offer curves illustrated in Figure 5 evaluate to the three different offers ϕ_0, ϕ_1 and ϕ_2. Since z, $\underline{z}$, and $\boldsymbol{r}$ are identical for each of these evaluations, the costs of provision must also be identical. All that changes between these offers is the level of profit accruing to the landlord. Thus the vertical distance between two offer curves (in which the arguments entering the cost function remain unchanged) measures the difference in profit associated with the two curves. Of course, this should be evident from Equation (18).

In the left hand panel of Figure 5, therefore, the vertical distance between the middle and upper offer curves is the difference in profits associated with the two curves i.e. $\pi_2 - \pi_1$.

Figure 5: The landlord's offer curves

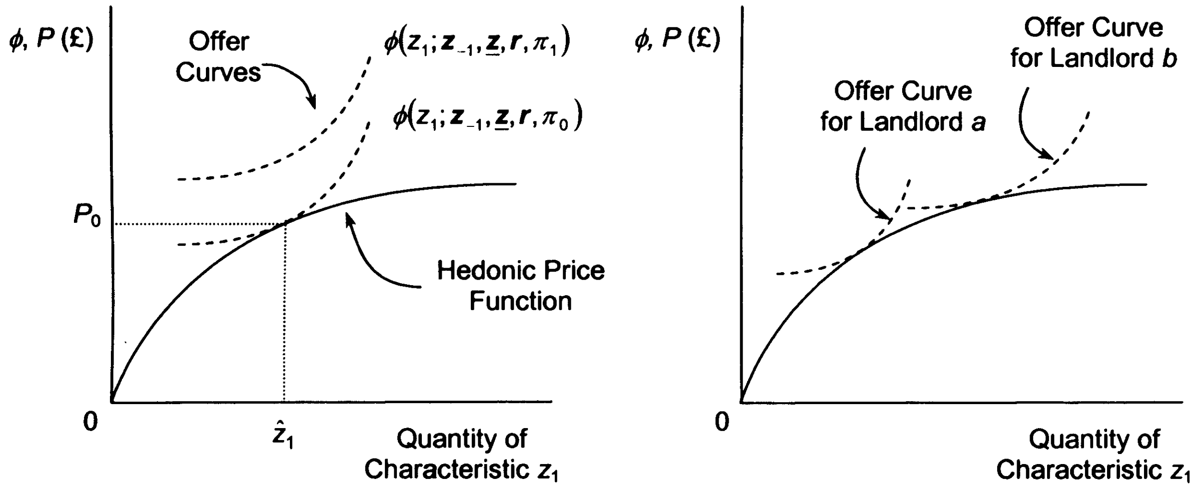


The right hand panel of Figure 5 depicts a comparable analysis, but this time we compare offer curves that differ in the level of the exogenously determined levels of $\underline{z}$. For example, we might imagine that the only difference between the upper and lower offer curves is the baseline level of crime in the area around the property. The upper curve is an offer curve for the landlord when faced by a low level of crime the, lower offer curve would represent the landlord's situation faced by a higher level of crime. Notice that the two offer curves are associated with the same level of profit to the landlord, π_1. At a particular level of provision of provision of z_1, again let us focus on z_1^*, these two curves evaluate to two different offers, ϕ^a and ϕ^b. Clearly, the change in crime has influenced the costs of the landlord in providing a property with attribute levels z_1^*, z_{-1}. For example, the lower the level of crime, the lower the level of vandalism, the less the landlord would have to spend in maintaining and repairing the property. Since the two offer curves represent the same profit to the landlord, the vertical distance between them must indicate this cost saving, i.e. $c^b - c^a$. Notice also that this cost saving does not necessarily have to result from costs incurred in the provision of attribute z_1.

In maximising profit, the landlord seeks to provide the bundle of housing attributes that positions them on the offer curve providing the highest level of profit whilst still being compatible with reigning market prices. Similar to the choices made by households, this entails a tangency condition. In the left hand panel of Figure 6, the highest offer curve compatible with the HPF is that returning a profit of π_0. The best this landlord can achieve, therefore, is to alter the level of attribute z_1 to $\hat{z}_1$, and charge a rent of P_0.

Again offer curves differ across landlords due to differences in purchase prices (though, for simplicity, this argument has been suppressed in the analysis), the vector of parameters $\boldsymbol{r}$ and the exogenously determined levels of attributes $\underline{\boldsymbol{z}}$. As a consequence, different landlords will choose to supply properties with different bundles of attributes. This is illustrated in the right hand panel of Figure 6. Indeed, the HPF forms the lower envelope to the set of all landlords' optimising offer curves.

Figure 6: Landlord's Optimising Choices of Housing Attributes to Supply

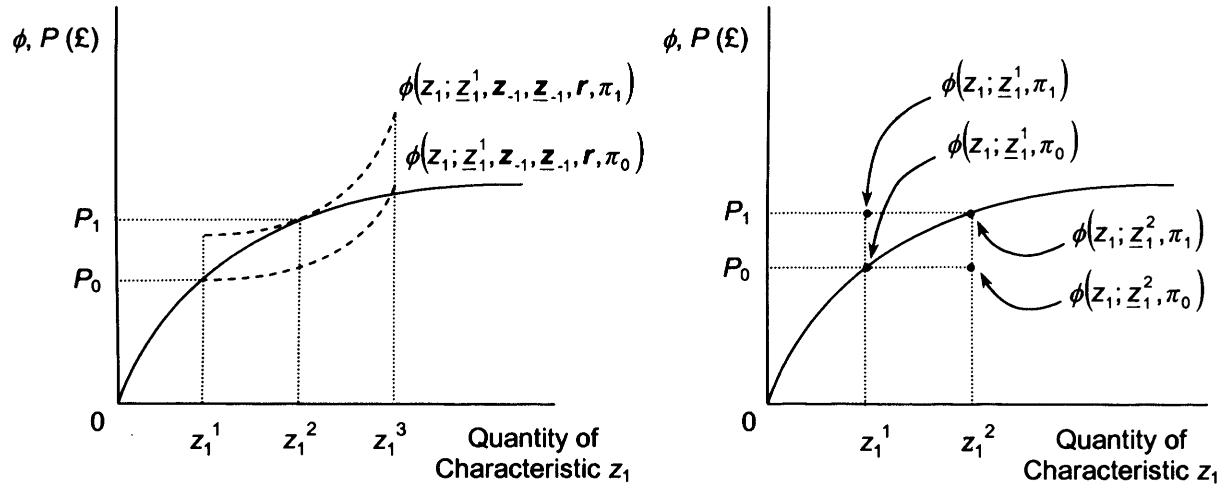


Offer curves differ slightly from their counterpart, bid curves, in so much as they will frequently be defined over quite a small range of attribute space. That is to say, for any one landlord the ability to change the attributes of the property may be relatively limited.

Let us use as an example the "peace and quiet" attribute of a property. To a large extent, this is fixed for the property by exogenous factors, most notably proximity to noisy roads. The left hand panel of Figure 7 shows offer curves for a landlord whose property is exposed to an exogenously determined level of peace and quiet, z_1^1. As such, this quantity enters the offer function where it is labelled $\underline{z}_1^1$. At this level of the attribute the landlord's property would command a rent of P_0 that would reward the landlord with a profit of π_0.

However, this is not the most profit that the landlord could earn on this property. The landlord could increase the level of peace and quiet at the property by, for example, planting trees in the front garden to act as a barrier to traffic noise and/or fitting sound-proofing windows. The best the landlord could do is to increase the peace and quiet attribute of the property to z_1^2. Here he would be able to a charge a higher rent, P_1, and would enjoy a profit of π_1.

Figure 7: Landlords' Optimising Choice when the level of provision of an attribute is constrained



In the example shown in the left hand panel of Figure 7, no amount of investment will increase the peace and quiet at this property beyond z_1^3, such that the offer function is not defined for values of z_1 in excess of this quantity. As it happens, these restrictions will not concern this landlord as he maximises his profit by providing a property with z_1^2 of the attribute.

However, the possibility arises that landlords may face a corner solution. To illustrate with our current example, the offer curve is constrained by the exogenously determined level of ambient noise pollution, which places a lower limit of z_1^1 on the peace and quiet of the property, and by the limitations of noise avoidance technology, which places an upper limit of z_1^3 on the peace and quiet of the property. If either of these limits were the profit maximising solution for the landlord then the simple tangential condition shown in Figure 6 would not hold.

In the extreme, landlords may have no control over the level of an attribute, in which case the lower and upper limits for an attribute are one and the same. Take for example the proximity of the property to local amenities such as a shopping centre or school. Since the property is fixed in space there is no way in which the landlord can influence the time it would take a household living in that property to access these facilities. In this case, the bid curves will shrink to a point above the exogenously determined level of the attribute.

Such a situation is illustrated in the right hand panel of Figure 7. Again to illustrate in one dimension we assume that the levels of all other property attributes do not change and for ease of exposition we suppress the vectors $\mathbf{z}_{-1}$, $\underline{\mathbf{z}}_{-1}$ and $\boldsymbol{r}$, from the offer function. Consider first the situation where the level of attribute z_1 is exogenously constrained to a level z_1^1. The best this landlord can do is to charge a rent of P_0, which returns a profit of π_0. Charging anything lower would necessitate missing out on possible profits, charging anything greater would make it impossible to rent out the property.

The only way in which this landlord could increase profits would be if there were an exogenous change in the level of z_1 enjoyed at the property. For example, if attribute z_1 represents accessibility to the town centre, then the building of an urban tram link that passed near the property would increase the accessibility of the property and consequently increase the level of attribute z_1. In Figure 7, this is represented by an increase from z_1^1 to z_1^2.

If we assume that this change doesn't influence the cost of providing other property attributes then the landlord could continue to charge P_0 and earn a profit of π_0. Of course, as illustrated in the figure, the landlord could make more profit than this by increasing the rent on the property to P^1. Charging this rent the landlord's profits increase to π_1.

We have dwelt on corner solutions such as those illustrated in the right hand panel of Figure 7 because such solutions typify environmental attributes and it is these attributes that are our central concern.[13] However, it is fair to assume that in the short run the levels of all attributes are constrained in a similar manner. For example, given the HPF, it may increase the profitability of a particular property if it were to possess an extra bedroom. Of course, such changes do not happen overnight. The landlord might have to employ an architect to design an extension to the property, apply for planning permission, employ builders and finally have the proposed extension constructed, decorated and furnished. We shall return to such considerations in the next chapter.

6. The Property Market: Equilibrium

So far we have examined the choices of consumers (households) and suppliers (landlords) in the property market independently. Figure 8 presents both sets of decisions combined in the same diagram. Households define their optimal residential location by choosing a property that boasts the set of attributes that coincide with the tangent of their lowest bid curve with the HPF. The household cannot increase their utility by bidding for a property with different characteristics. Simultaneously, landlords maximise their profits by choosing to supply a property with the set of attributes that allows them to move to their highest offer curve that is still compatible with market prices. Supplying an alternative set of attributes would result in the landlord receiving offers for the property that resulted in lower levels of profits.

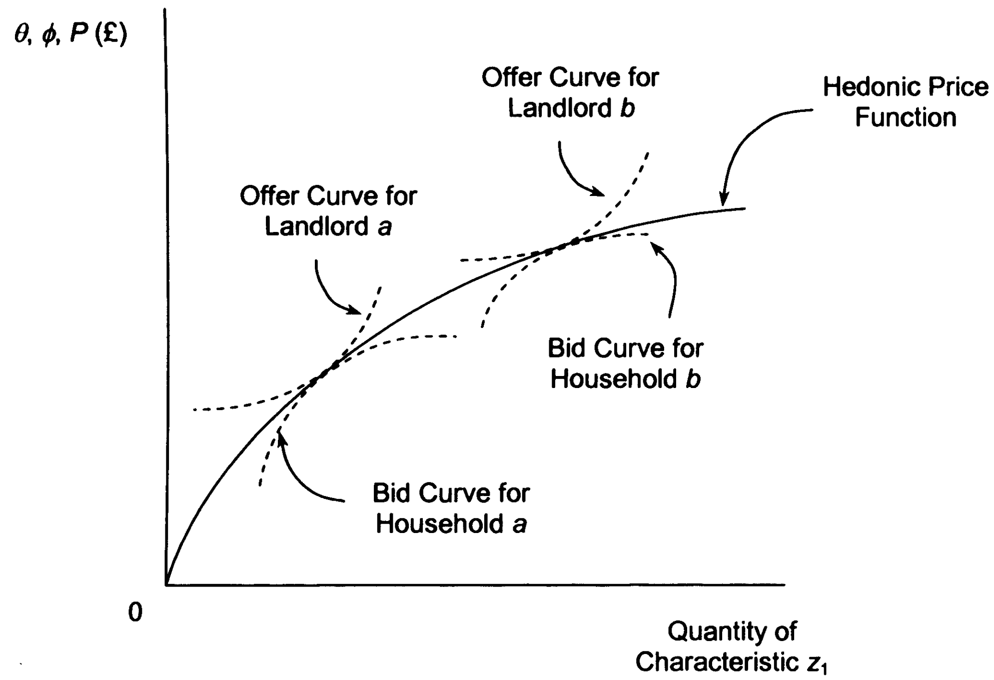
As depicted in Figure 8, the bid curves of households and the offer curves of landlords will "kiss" along the HPF. At each coincidence of bid and offer curves, a landlord and household are paired; the landlord can do no better than to accept the household's offer who in turn can do no better than to rent the property from that landlord.

---

[13] Since many environmental qualities have the properties of public goods, their level of provision tends, to a greater extent, to be determined exogenously. Further, landlords frequently have limited scope for adjusting the levels of these attributes through investment in private goods for the property.

The situation we have described is one of market equilibrium. At the reigning market prices revealed by the hedonic price schedule, demand for properties is equal to the supply of properties and the market clears. Hitherto our analysis has been at the level of individual households and landlords. At this disaggregate level we have assumed that each individual economic agent, being only a small player in the entire market, takes the equilibrium hedonic price schedule as given. To understand how the equilibrium is reached in the first place, we need to look at how households and landlords interact in the aggregate.

Figure 8: Choice of Property Attributes for Different Households



7. Equilibrium: A general solution

Thus far, we have attributed differences in the bid curves of households and differences in the offer curves of landlords to variation in the characteristics of those households and landlords (represented by the vectors s and r respectively). To build a more formal model of equilibrium in hedonic markets we need to make more concrete statements concerning the distribution of household preferences determined by s and landlord production costs determined by r. Loosely following, Ekeland et al. (2002), let us define household preferences as being determined by two sets of parameters;

- $\boldsymbol{A}$ which represents preference parameters held commonly by all households
- $\boldsymbol{\alpha}$ which represents preference parameters that vary across households (due to differences in $\boldsymbol{s}$).

As such we can rewrite the utility function as;

$$U(z, x; \boldsymbol{A}, \boldsymbol{\alpha}) \tag{5a}$$

Similarly, we can represent landlord production costs by the two sets of parameters;

- $\boldsymbol{B}$ which represents cost parameters held commonly by all landlords
- $\boldsymbol{\beta}$ which represents cost parameters that vary across landlords (due to differences in $\boldsymbol{r}$).

As such we can rewrite the profit function as;

$$C(z; \underline{z}, \boldsymbol{B}, \boldsymbol{\beta}) \tag{16a}$$

Where, to reiterate, $\underline{z}$ is the level of the attributes that, following initial acquisition of the property, are supplied costlessly to the landlord. For simplicity, we assume here that $\underline{z}$ is identical for all landlords.

Finally, let us assume that the HPF takes a form that is determined by two sets of parameters; $\boldsymbol{\gamma}$ and $\boldsymbol{\Gamma}$, giving us;

$$P(z; \boldsymbol{\Gamma}, \boldsymbol{\gamma}) \tag{2a}$$

Our model of the property market aims to discover the values of $\boldsymbol{\gamma}$ and $\boldsymbol{\Gamma}$, that result in a market equilibrium.

Now the household's problem is to maximise their utility (Equation 5a) subject to the constraint imposed by their budget (Equation 6) requiring $x = y - P(z; \boldsymbol{\Gamma}, \boldsymbol{\gamma})$. In the same way the landlord seeks to maximise profits, now given by;

$$\pi(z; \underline{z}, \boldsymbol{B}, \boldsymbol{\beta}, \boldsymbol{\Gamma}, \boldsymbol{\gamma}) = P(z; \boldsymbol{\Gamma}, \boldsymbol{\gamma}) - C(z; \underline{z}, \boldsymbol{B}, \boldsymbol{\beta}) \tag{17a}$$

In addition to the budget constraint, the first order conditions for a maximum are obtained by differentiating the utility function (Equation 5a) and profit function (Equation 17a) with respect to the choice of property characteristics *z* and setting the resulting expression to zero:

$$U_z(z,x;\boldsymbol{A},\boldsymbol{\alpha}) - p_z(z;\gamma,\boldsymbol{\Gamma})\,U_x(z,x;\boldsymbol{A},\boldsymbol{\alpha}) = \mathbf{0} \tag{19}$$

$$p_z(z;\boldsymbol{\Gamma},\gamma) - c_z(z;\underline{z},\boldsymbol{B},\boldsymbol{\beta}) = \mathbf{0} \tag{20}$$

where;

- U_z is the vector of partial derivatives of the utility function with respect to the property characteristics *z*; the marginal utilities of the characteristics.
- U_x is the partial derivative of the utility function with respect to the composite good, the marginal utility of income.
- p_z is the vector of partial derivatives of the HPF with respect to the property characteristics *z*; the implicit prices.
- c_z is the vector of partial derivatives of the cost function with respect to the property characteristics *z*; the marginal costs of production.

Likewise, the second order conditions defining a maximum are given by;

$$\left(U_x p_{zz'} + p_z U_{xx'} p_z' - U_{zz'}\right) \text{ is negative definite} \tag{21}$$

$$\left(p_{zz'} - c_{zz'}\right) \text{ is negative definite} \tag{22}$$

To make things simple, it is assumed that the HPF, utility function and cost function are twice differentiable and that the second order conditions hold as strict inequalities.

Solving the FOCs (comprising the budget constraint, Equation (19) and Equation (20)) for *z* defines the *ordinary demand and supply functions*;

$$z^d = z^d(A,\alpha,\Gamma,\gamma,y) \tag{23}$$

$$z^s = z^s(\underline{z},B,\beta,\Gamma,\gamma) \tag{24}$$

where z^d is the vector of property characteristics demanded by a household given their income and preference parameters and the parameters of the HPF. Similarly, z^s is the vector of property characteristics supplied by a landlord given their cost parameters and the parameters of the HPF.

Notice that because marginal prices are non-parametric and hence are themselves functions of the quantities of property characteristics, it is not the implicit prices (p_z) that enter the demand and supply functions but the parameters of the HPF.

Assuming that a local implicit theorem holds, (23) and (24) can be inverted to isolate α and β the preference and cost parameters that vary across households and landlords respectively. Let us denote these functions;

$$\alpha = \alpha(z,A,\Gamma,\gamma,y) \tag{25}$$

$$\beta = \beta(z,\underline{z},B,\Gamma,\gamma) \tag{26}$$

Now, let the distribution of preference parameters across households be given by the multivariate density function f_α. Likewise let the distribution of cost parameters across landlords be given by the multivariate density function f_β. Since the demand and supply relationships are functions of these distribution functions, the change of variables technique can be used to determine the densities of demand and supply;

$$f_{z^d} = f_\alpha\left(\alpha(z,A,\Gamma,\gamma,y)\right)\left|\frac{\partial\alpha(z,A,\Gamma,\gamma,y)}{\partial z}\right| \tag{27}$$

$$f_{z^s} = f_\beta\left(\beta(z,\underline{z},B,\Gamma,\gamma)\right)\left|\frac{\partial\beta(z,\underline{z},B,\Gamma,\gamma)}{\partial z}\right| \tag{28}$$

Where;

- f_{z^d} gives the demand densities, that is, the quantity of households demanding each particular vector of property characteristics given the parameters of the HPF, preference parameters $\boldsymbol{A}$ and the distribution of $\boldsymbol{\alpha}$.

- f_{z^s} gives the supply densities, that is, the quantity of landlords supplying each particular vector of property characteristics given the levels of exogenously supplied characteristics, $\underline{z}$, the parameters of the HPF, the cost parameters $\boldsymbol{B}$ and the distribution of $\boldsymbol{\beta}$.

Equilibrium in the hedonic market requires that demand is equal to supply for all combinations of property characteristics. The market clearing HPF, therefore, will be defined by the parameters γ and Γ that solve the second order differential equation;

$$f_{\alpha}\left(\alpha(z,A,\Gamma,\gamma,y)\right)\left|\frac{\partial\alpha(z,A,\Gamma,\gamma,y)}{\partial z}\right| = f_{\beta}\left(\beta(z,\underline{z},B,\Gamma,\gamma)\right)\left|\frac{\partial\beta(z,\underline{z},B,\Gamma,\gamma)}{\partial z}\right| \quad (29)$$

a not inconsiderable task!

8. Equilibrium: Analytical and Simulated Solutions

Whilst the theoretical framework is quite general, the complexity inherent in solving Equation (29) explains why attention tends to have focused on a particularly simple specification of the model (examined by Tinbergen, 1956, Epple, 1987 and Tauchen and Witte, 2001, amongst others). Ekeland et al (2002) call this the Normal Linear Quadratic (NLQ) model since it assumes that the heterogeneity of households and landlords is *normally* distributed in the population, imposes *linear* demand and supply functions and thereby provides a closed-form solution in the shape of a simple *quadratic* equilibrium HPF with linear implicit prices. Heckman et al. (2003) show that in equilibrium the NLQ model predicts that a range of properties exhibiting different combinations of attributes will be provided to the market in equilibrium. Moreover, the density of these properties in attribute space is found to follow a normal distribution. We review the NLQ model in detail in Section A1 of Appendix A.

In a series of recent papers, Ivar Ekeland, James Heckman, Rosa Matzkin and Lars Nesheim (Ekeland et al., 2002, 2003; Heckman et al., 2002, 2003) have investigated the nature of the market equilibrium when some of the restrictive assumptions of the NLQ model are relaxed. Again a detailed review of these models is consigned to Section A2 of Appendix A. Since, these problems no longer provide closed-form solutions, numerical methods are used to approximate the HPF and characterise the market in equilibrium. Their analysis reveals that even minor perturbations from the assumptions of the NLQ model disrupt the neat simplicity of the equilibrium solution.

For example, Ekeland et al. (2003) abandon the assumption of normally distributed heterogeneity within the populations of households and landlords. Instead they model heterogeneity as the mixture of two different normal distributions. Naturally, the greater the degree of mixing, the further the distribution of heterogeneity strays from normal. Ekeland et al. (2003) discover that when the distribution of heterogeneity is non-normal, the market equilibrium is no longer characterised by a quadratic HPF with linear implicit prices. Rather, the greater the degree of mixing, the greater the degree of non-linearity observed in the implicit price functions. Imposing economically reasonable restrictions (i.e. positive implicit prices, only positive quantities of attributes demanded and supplied) only serves to exaggerate the nonlinearity of implicit prices. Indeed as Ekeland et al. (2003) prove in the context of the NLQ model, the implicit price schedules of the equilibrium HPF are generically nonlinear.

Heckman et al. (2003) also examine the equilibrium density of properties exhibiting different levels of attributes. In the NLQ model, this density follows a normal distribution. They observe that as the distribution of heterogeneity is made increasingly non-normal, the density of properties in attribute space follows suit.

Heckman et al. (2003) extend their investigation by examining models in which the quadratic specifications of the household utility and landlord cost functions are replaced with higher order polynomials. Again, the increased flexibility of these specifications precipitates increasing nonlinearity in the implicit price schedules of the equilibrium HPF. What is more, in these more flexible models, the equilibrium density of properties in attribute space is far from normally distributed. Indeed, in the cases illustrated in Heckman et al. (2003) the density of supply exhibits many modes; that is, in equilibrium there exists clusters of properties exhibiting similar combinations of attributes, whilst properties with other combinations of attributes are sparsely represented in the equilibrium market. As Heckman et al (2003) point out, "the model is capable of generating equilibria in which there are nearly gaps in the range of products marketed".

9. Equilibrium: Interaction-Based Models

In a parallel theoretical literature. Nesheim (2002) investigates the nature of property market equilibria in which households choose where to live based on their willingness to pay (WTP) for locational quality. In particular, he concerns himself with neighbourhood effects. That is, a model where households' valuations of properties depend not on the characteristics of the properties themselves but on the characteristics of the equilibrium sets of people that choose to inhabit the neighbourhood in which that property is located. Nesheim's model is reviewed in detail in Section A3 of Appendix A.

Paralleling the work of Ekeland et al. (2003), Nesheim finds that in all but the simplest cases, the curvature of the equilibrium hedonic price schedule is highly nonlinear. Indeed, Nesheim reports that for certain parameter values, a kinked price function is required in order to attain an equilibrium.

Similarly, Nesheim (2002) finds that in equilibrium the property market may be characterised by lumpy provision. Neighbourhoods boasting high and low levels of quality are relatively more common than those at intermediate levels. Moreover, Nesheim (2002) shows that households will sort themselves across the urban area such that the traits of households within a neighbourhood are likely to be less varied than those of the population as a whole. Indeed, the more correlated a trait is with WTP for locational quality, the more homogenous neighbourhoods are likely to be in this trait and the greater will be the differences in the average level of this trait across neighbourhoods. As such, the interaction-based models predict that households may sort according to their personal socioeconomic characteristics. Indeed, in the equilibrium market, neighbourhoods with residents showing particular combinations of characteristics may be well-represented, whilst neighbourhoods with residents exhibiting other combinations may be relatively rare.

10. Summary and Conclusions

In this chapter we have laid out the fundamental economic theory that underpins hedonic analysis of property markets. The basic model as developed by Rosen (1974) describes how the interactions of households and landlords in the property market result in an equilibrium maintained by an equilibrium HPF.

In this model households want good quality residences at low prices whilst landlords wish to let their property at high prices so as to make the most profit possible. The market reconciles these conflicting goals through the price mechanism. An equilibrium is attained when the hedonic price schedule ensures households (within their limited budgets) cannot increase their utility by choosing a different property and landlords cannot increase their profits by increasing the property's rent or changing its characteristics.

We presented two particular versions of the model. One in which households value a property based on the characteristics of the property itself, one in which they value a property based on the equilibrium set of people living in the neighbourhood of that property. Both these models provide valuable insights into the nature of the equilibrium HPF.

In particular, the equilibrium hedonic price schedule is found to be a complex function of households' preferences, landlords' costs, exogenously determined property attributes and how these are distributed across households and landlords in the urban area. Changes in any of the parameters of the model cause shifts in the equilibrium hedonic price schedule. In particular, exogenous changes in the characteristics of properties, such as those that might result from government interventions to improve environmental quality, have direct but complex implications for the equilibrium HPF.

Apart from this fairly nebulous set of conclusions, several more testable hypotheses emanate from the models. First, in all but the simplest cases, the models predict that the equilibrium price schedule will be highly nonlinear. Indeed, Nesheim (2002) finds that for certain parameter values, a kinked price function is required in order to attain an equilibrium. Second, both models predict that in equilibrium the market may be characterised by "lumpy" supply. For example, in a number of the models investigated by Heckman et al. (2003), equilibrium is characterised by a dense supply of properties providing certain combinations of property characteristics and a relative paucity of properties providing other combinations.

If real world property markets are behaving in ways suggested by these models, then we would expect to see well-defined clusters of properties providing similar combinations of characteristics and/or neighbourhoods inhabited by households with relatively homogenous characteristics. If such clusters exist then, by definition, the properties within them will be more closely located on the hedonic price surface. In Part 2 of this thesis we explore how this observation might be exploited in specifying an econometric model for the equilibrium HPF.

CHAPTER 2. WELFARE MEASUREMENT IN HEDONIC PROPERTY MARKETS

1. Introduction

Our interest in hedonic markets stems from the fact that environmental quality can be counted amongst the attributes of a property. Whilst the various attributes which make up environmental quality (e.g. peace and quiet, clean air, access to recreational areas etc.) are frequently not directly traded in markets, hedonic property markets provide an indirect means by which households can express preferences for such goods.

Like the other attributes of a property, differences in environmental quality will be reflected in differences in the price paid for a property. Indeed, with information on the implicit price of environmental quality and the residential locations chosen by different households, analysts have access to information from which they can deduce household preferences for environmental goods.

The search for these underlying preferences is the key goal of empirical analysis of hedonic market data. Specifically, establishing the structure of preferences makes it possible to estimate the impact on economic welfare of a change in environmental quality. We shall return to the issue of estimating household preferences from empirical data in the final part of this thesis.

In this chapter we show how the theory of hedonic markets outlined in Chapter 1 allows us to describe the welfare effects of changes in environmental quality.

2. The Hedonic Market and Changes in Environmental Quality

Before we embark on an analysis of welfare measures it is worth developing an intuitive understanding of the impact that a change in environmental quality might have in the property market. Let us consider the impacts of an environmental improvement such as a reduction in noise pollution, a fall in levels of crime or an increase in air quality. This change may be a relatively minor or a significant environmental improvement. Likewise, the improvement may take place uniformly across the urban area or be restricted to specific neighbourhoods.

Marginal, localised changes may have little impact on the property market as a whole. Of course landlords will be able to increase the rent they charge on properties in the improved area since those properties now enjoy a higher level of environmental quality. Household's living in these properties will find themselves paying more than previously. Indeed, to return to an optimal residential location they may elect to move to a new house possessing the original bundle of characteristics charged at their previous rent. In the real world, however, there are considerable transaction costs associated with moving house. For relatively small changes in rent, therefore, households may elect to remain where they are. Nevertheless, in the long run, perhaps prompted by other changes in the property market, we would expect households to move to a property with an optimal bundle of characteristics.

If the environmental change is neither marginal nor localised then the pattern of changes in the property market may be far more complex. As for any good, significant changes in the conditions of supply and demand may lead to changes in market-clearing prices. Since the geographical nature of property markets ensures they are relatively small (compared, say, to goods traded in world markets) we might expect them to be particularly responsive to relatively minor changes in market conditions.

Bartik (1988) provides a detailed description of how an environmental improvement might impact on property rents, location choices and housing supply. The complex pattern of readjustments in the market involve a tendency for rental prices to increase in the area enjoying the environmental improvement counteracted by a general reduction in the price of environmental quality across the entire market required to ensure market clearance. Simultaneously, demand for property characteristics that are substitutes for the environmental attribute will decline. For instance, demand for double-glazed properties will decline in an area in which noise pollution has been reduced. Similarly, demand for complementary attributes will increase. For example, a reduction in air pollution might increase demand for houses with gardens. The implicit prices for these substitutes and complements will themselves have to adjust in order to ensure that the demand for these attributes is balanced by the supply.

Further, in response to the shifts in the HPF, households may choose to relocate. Indeed, changes in the socioeconomic characteristics of households in different neighbourhoods may encourage landlords to adjust their properties' characteristics in order to maximise their profits.

It is evident that the overall change in the HPF and the resulting changes in rents and locational choices are complex. For any one property, the eventual rental value will not be determined solely by the change in environmental quality experienced at that location. Instead it will be determined by the complex interaction of supply and demand across the entire market.

3. Measuring Changes in Economic Welfare in Hedonic Markets

Our goal in analysing data from hedonic markets is to establish how changes in environmental quality impact upon economic welfare. That is, we are seeking to measure how greatly changes in environmental quality change the well-being of economic agents in the property market. We have defined these economic agents as *households* and *landlords*. Further, we have defined household well-being as the *utility* they derive from their choice of residential location and expenditure on other goods, whilst landlord well-being is defined as the *profits* they realise from rental of their property.

For landlords then, the effect on economic welfare resulting from a change in environmental quality can be measured as the change in their profits ($\Delta\pi$) from renting out a property.

For households, the ideal measure of change in economic welfare would be to measure the change in utility (ΔU) that the household experienced as a result of the change in environmental quality. Of course, we have no way of deriving such a measure. Rather we focus on a monetary valuation of ΔU known as the *compensating measure* of welfare change. Compensating measures take the current level of household utility as a baseline.

- For an environmental improvement, the compensating measure would be the maximum quantity of money that the household would willingly give up in order to ensure that they enjoyed the environmental improvement; the household's maximum willingness to pay (*WTP*) to achieve an improvement.

- For a fall in environmental quality, the compensating measure would be the minimum amount of money that the household would accept in order to endure the deterioration in environmental quality; the household's minimum willingness to accept (*WTA*) compensation for a deterioration

With these measures in mind, let us consider a property market and examine the welfare impacts of a change in environmental quality. As shall become apparent in the following two sections, this is not as straightforward as might be hoped. It turns out that there are a number of ways in which the change in economic welfare might be evaluated; each evaluation differing according to the assumptions that are made concerning the response of households and landlords to the change in environmental quality. As might be expected, the fewer the assumptions that are made, the more comprehensive the measure of the welfare change. At the same time, unfortunately, the fewer the assumptions that are made the greater the informational requirements involved in calculating the welfare measure.

4. Changes in Economic Welfare for Households

Let us assume that the property market we are considering is in equilibrium. In this market (following Bartik's, 1988, notation) we shall denote the original equilibrium HPF by $P^b(z)$, where the superscript *b* indicates that this is *before* any changes in conditions in the hedonic market. Following a change in environmental quality, the market settles at a new HPF that we shall denote $P^a(z)$. Once again the superscript *a* indicates that this is *after* the change in conditions in the hedonic market.

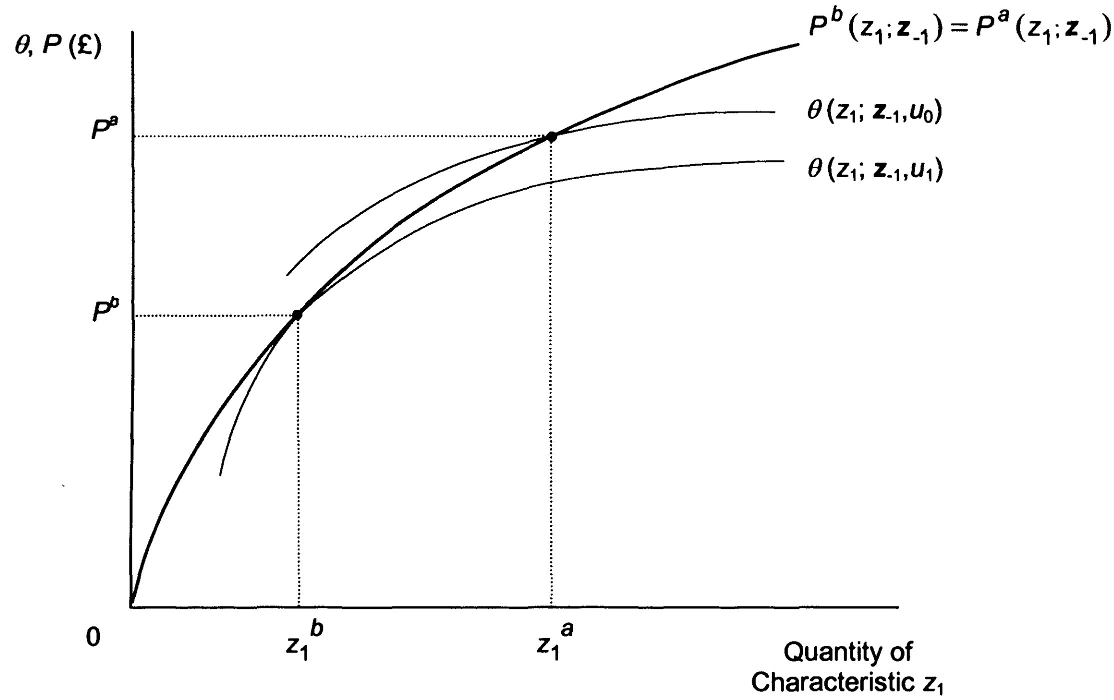
4.i. Welfare changes from a localised environmental improvement

To begin with let us consider the welfare impact of a localised environmental improvement. We assume that such a change will not impact on the property market as a whole and the HPF will not need to adjust in order to clear the market. In this case, $P^a(z)=P^b(z)$.

Let us focus on a property located in the area experiencing the environmental improvement. If we designate attribute z_1 to be the level of environmental quality, then the initial level of environmental quality at this property can be represented by z_1^b. As illustrated in Figure 1, at this level of environmental quality the property commands a rental price of P^b.

The household choosing to reside in this property will have a bid curve tangential to the HPF at this level of environmental quality. In Figure 1, this utility maximising choice places the household on their lowest bid curve compatible with the HPF and results in a level of utility u_1.

Figure 1: Change in household welfare from a localised change in environmental quality and costless moving



Now, the exogenous improvement increases the environmental quality of the property from z_1^b to z_1^a (where once again *b* superscript stands for *before* and *a* superscript stands for *after*). Of course, improving the attributes of the property enables the landlord to charge a higher rental price. Indeed, the rent on the property following the environmental improvement would increase from P^b to P^a. Clearly, this represents a benefit to the landlord but we shall postpone a discussion of this welfare gain until the next section.

What then are the welfare impacts on the household residing at this location? The household enjoys an improvement in environmental quality but this is accompanied by an increase in rental price. As illustrated in Figure 1, the household will find itself at a less than optimal residential location. Indeed, continuing to live at the property would mean their level of utility would fall from u_1 to u_0.

Since the HPF has not changed, we know that the household's optimal choice would be an identical property that boasted the original level of environmental quality. Indeed, if we assume that moving house is costless, then the household would be best off moving back to just such a property. By so doing they would return to their original level of utility, u_1. As such, under the assumption of costless moving, the environmental improvement should have no impact on the households' welfare.

In the real world, however, moving properties is not costless. Contrary to the argument presented by Palmquist (1991, 1992), however, not the entirety of these transactions costs should be considered as welfare losses. In particular, over time households' residential preferences will change (e.g. as a result of marriage, the birth of children, retirement etc.). Foreseeing such changes, we must assume that households plan to move properties at certain times in their life, independent of market conditions, and anticipate the transactions costs they will incur in so doing. It is only the extent to which changes in the property market precipitate households to move unexpectedly or in advance of anticipated property moves that should be considered. Let us label this added cost translated into an equivalent per period amount, *tc*.

Following the localised environmental change, two possible responses are open to households;

- If the benefits of moving outweigh the transactions costs then the household will relocate to a new property with the attributes of their original choice. In welfare terms, the household ends up enjoying the same level of utility as prior to the environmental change but are worse off by an amount equal to the costs of moving. Thus the quantity *tc*, measures the per period welfare loss of the environmental improvement.

- Alternatively, if the benefits of moving do not exceed the transaction costs then the household will decide to remain in their original, though now sub-optimal, residential location. Consequently, *tc* gives an upper bound on the welfare loss to the household.

When considering transactions costs in the context of localised environmental change, we can state that a household's loss in welfare cannot exceed *tc* since they could always pay this amount so as to relocate to a property offering the level of welfare enjoyed prior to the change.

If the total number of properties in the market is labelled *H* then the small subset of properties affected by the environmental improvement can be labelled H_1. Further, if we index all the households in the market by $h = 1$ to *H*, then the welfare change experienced by households from a localised environmental improvement can be expressed;

$$\sum_{h \in H_1} -tc_h \leq W_H \leq 0 \tag{1}$$

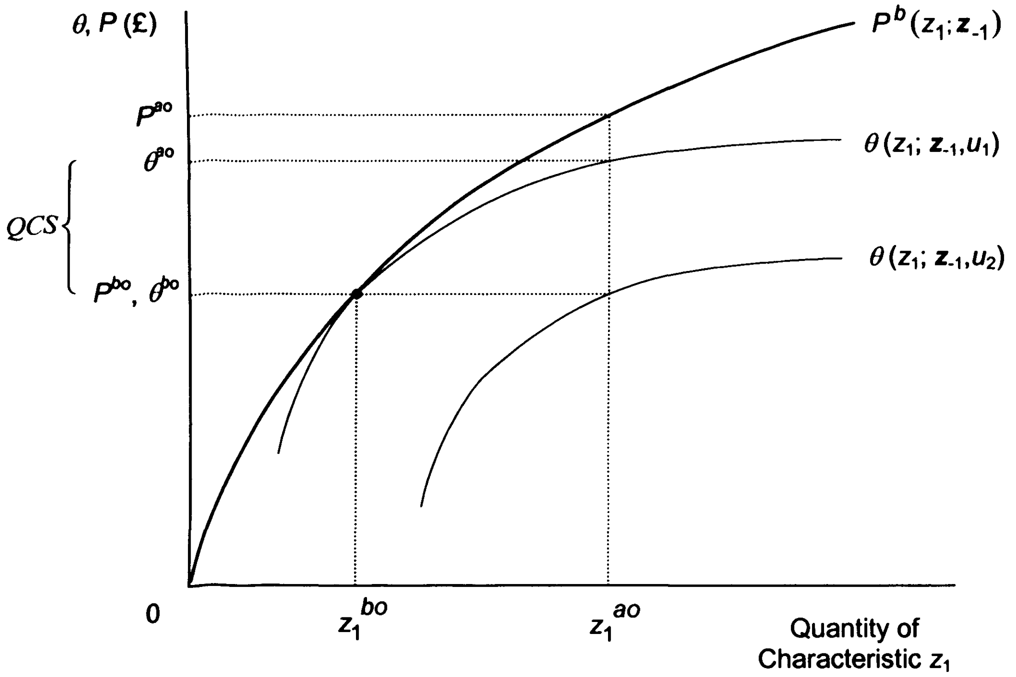
where W_H is the total welfare change experienced by households in the market.

4.ii. Welfare changes from a non-localised environmental improvement

Imagine now, that we are dealing with an environmental improvement that has more than a purely localised impact. If the change we are considering represents a major improvement and/or is widely spread across the urban area then the consequences for the property market may extend beyond a simple increase in the price of affected properties.

Figure 2 presents the situation facing a household living in the area witnessing an environmental improvement. At the original level of environmental quality the household faces the old hedonic price schedule, $P^b(z)$, and maximises its utility by choosing a property with a level of environmental qualtiy indicated by z_1^{bo}. Here we have expanded the notation such that the superscript *bo* indicates that this is the quantity chosen *before* the change in environmental quality in the households *old* choice of property. At this point, the household reaches its lowest bid curve that is still compatible with the prices it faces in the market, $\theta(z;u_1)$.[14] The household's WTP or bid, indicated by θ^{bo}, is equal to the market price, P^{bo}, and the household enjoys a level of utility labelled u_1.

Figure 2: The Quantity Compensating Surplus measure of the welfare change resulting from an improvement in environmental quality



An exogenous change increases the environmental quality enjoyed at this location to z_1^{ao}, where the superscript *ao* indicates that this is the environmental quality *after* the change in the household's *old* choice of property.

Since we are now considering a non-localised change we assume that the HPF shifts in response to this environmental improvement. However, for the moment, we shall ignore this general equilibrium response. Further we shall consider the situation in which landlords continue to charge the rental price associated with old level of

[14] Income and socioeconomic characteristics have been suppressed to simplify notation.

provision of z_1. In this case, the household has effectively been given the benefits that come from living in a location with an improved environment. Indeed, at this location paying the original level of rent for that property they would find themselves on the bid curve $\theta(z; u_2)$ realising a higher level of utility labelled u_2.

The household's WTP to enjoy this benefit is illustrated in Figure 2. The bid curve on the diagram traces out all combinations of WTP and levels of environmental quality that result in the household enjoying a level of utility labelled u_1. Of course this is also the level of utility that the household realised prior to the environmental change by locating at their optimal residential location. To achieve this level of utility the household was willing to pay θ^{bo}. Following the environmental improvement the household would be willing to pay θ^{ao} to achieve the same level of utility. A measure of the household's WTP for the change in environmental quality is the difference between these two amounts.

This amount has been termed the *quantity compensating variation*, by Palmquist (1988). However, following Freeman's definitions (see Freeman, 1993; p 48-9) this is probably best thought of as a *compensating surplus* measure since it allows for no adjustment in household residential location following the change in environmental quality. Hence here we label this amount as the *quantity compensating surplus* (*QCS*). This amount is shown graphically in Figure 2 and can be stated mathematically as;

$$QCS = \theta\left(z_1^{ao},\ z_{-1}^{bo}; y, s, u_1\right) - \theta\left(z_1^{bo},\ z_{-1}^{bo}; y,\ s, u_1\right) \tag{2}$$

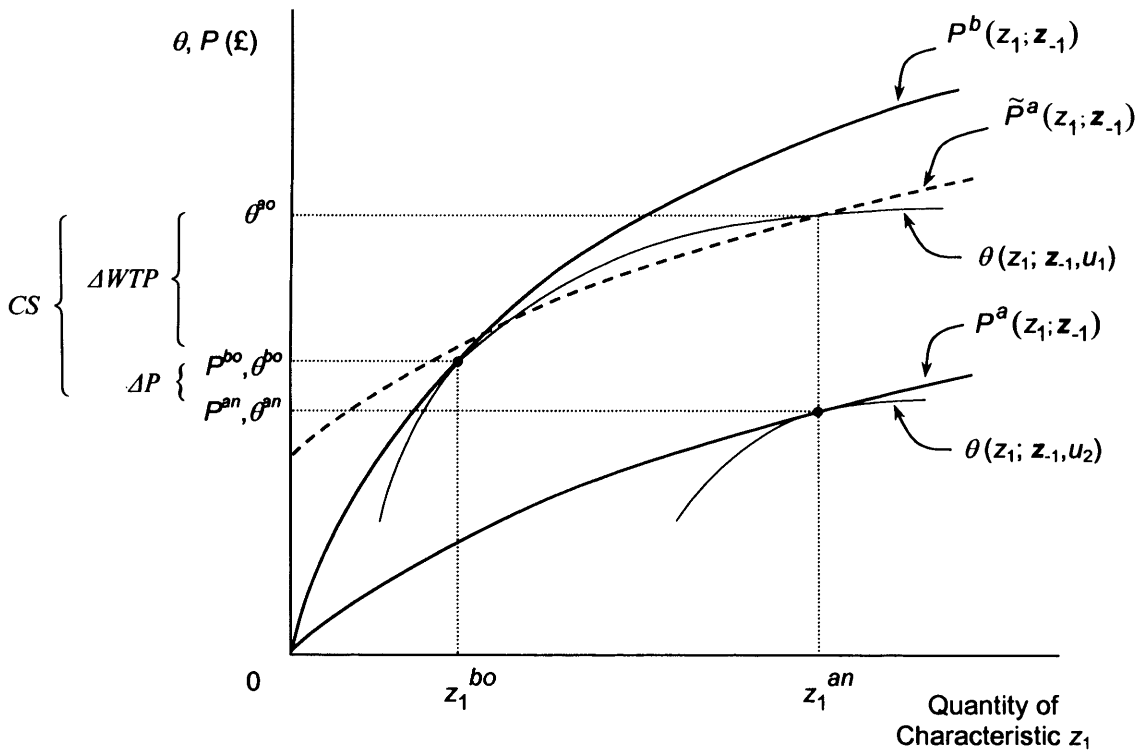
Since, the *QCS* measure assumes there are no adjustments in the hedonic property market the welfare change is assumed to impact only households in the affected area. Indeed, using this measure, the total welfare impact of the environmental improvement is given by;

$$W_H = \sum_{h \in H_1} QCS = \sum_{h \in H_1} \theta\left(z_{1h}^{ao},\ z_{-1h}^{bo}; y, s, u_{1h}\right) - \theta\left(z_{1h}^{bo},\ z_{-1h}^{bo}; y, s, u_{1h}\right) \tag{3}$$

Notice that the informational requirements of the *QCS* measure are relatively undemanding. To evaluate W_H using this measure, a researcher would simply have to know the environmental quality at all affected properties before and after the improvement and be able to evaluate the bid function for each household at these two values of environmental quality.

However, the *QCS* measure of welfare change is relatively restrictive in the assumptions it makes concerning how the market and the economic agents in the market react to a change in environmental quality. For example, as illustrated in Figure 3, large scale improvements in environmental quality in the urban area will change the conditions of supply in the market. Since environmental quality is relatively more abundant, the HPF may shift so as to reduce the price of property at any given level of environmental quality. The hedonic function *after* this adjustment is labelled $P^a(z)$.

Figure 3: The Compensating Surplus measure of the welfare change resulting from an improvement in environmental quality



Since the hedonic function has changed the environmental improvement will impact all households in the property market. As a consequence we would expect each household in the property market to adjust to the new HPF by choosing a new residential location. For example, the household illustrated in Figure 3, maximises their utility by moving to a new property with a level of environmental quality given by z_1^{an}. Where the superscript *an* indicates that this is the level of environmental quality enjoyed *after* the change at their *new* choice of residential location. By moving property, the household moves on to a lower bid curve and manages to achieve a higher level of utility, u_2.

Since, households are allowed to respond optimally to the changes in the hedonic market by moving residential location, our previous measure of welfare change, the *QCS,* is no longer an adequate measure of the benefits of an environmental improvement. In particular, the household in Figure 3, finds itself enjoying a level of utility, u_2, that exceeds that enjoyed prior to the environmental improvement in their original location, u_1. We assume that they would be willing to pay to maintain this improved well-being. Indeed, the monetary measure we seek is the amount of money that once taken away from the household in their new residential location would return them to their original level of well-being.

Figure 3 can be used to illustrate this second measure of welfare change that allows for household relocation. Here the change in the household's income that would result from paying out a compensating monetary measure, is shown as a vertical shift in the HPF. In effect, paying out money is equivalent to making all properties more expensive[15]. The maximum amount the household would be willing to pay to ensure the change in environmental quality whilst constrained to remain at their new residential location, will be the amount that shifts the HPF to the point where it intersects the original bid curve.

As illustrated in Figure 3, the vertical distance between the hedonic function $P^a(z)$ and the hedonic function as it would appear to the household once it had paid out its maximum WTP, $\widetilde{P}^a(z)$, gives a second measure of welfare change. This distance is

---

[15] Readers familiar with the illustration of welfare measures in diagrams with indifference curves and income constraints will recognise this procedure. Indeed, this parallel is made explicit by remembering that the bid curve and HPF are simply inversions of corresponding indifference curves and income constraints (see Chapter 1).

the *compensating surplus* (*CS*) measure of the household's welfare change described by Bartik (1988).

This *CS* measure can be decomposed into two separate values. The first value is the household's WTP for the change in housing attributes. That is, the difference between the household's WTP to achieve a level of well-being u_0, at the old and new residential locations (*ΔWTP*). The second value is simply the difference in rental payments at the old and new residential locations (*ΔP*). In mathematical terms, therefore, *CS* can be written as;

$$CS = \Delta WTP - \Delta P = \begin{pmatrix} \theta\left(z_1^{an}, z_{-1}^{an}; u_1\right) - \theta\left(z_1^{bo}, z_{-1}^{bo}; u_1\right) \\ -\left[P^a\left(z_1^{an}, z_{-1}^{an}\right) - P^b\left(z_1^{bo}, z_{-1}^{bo}\right)\right] \end{pmatrix} \tag{4}$$

Since all households are assumed to relocate in response to the shift in the HPF the total welfare benefits of the environmental improvement will include a measure for each of the *H* households in the urban area;

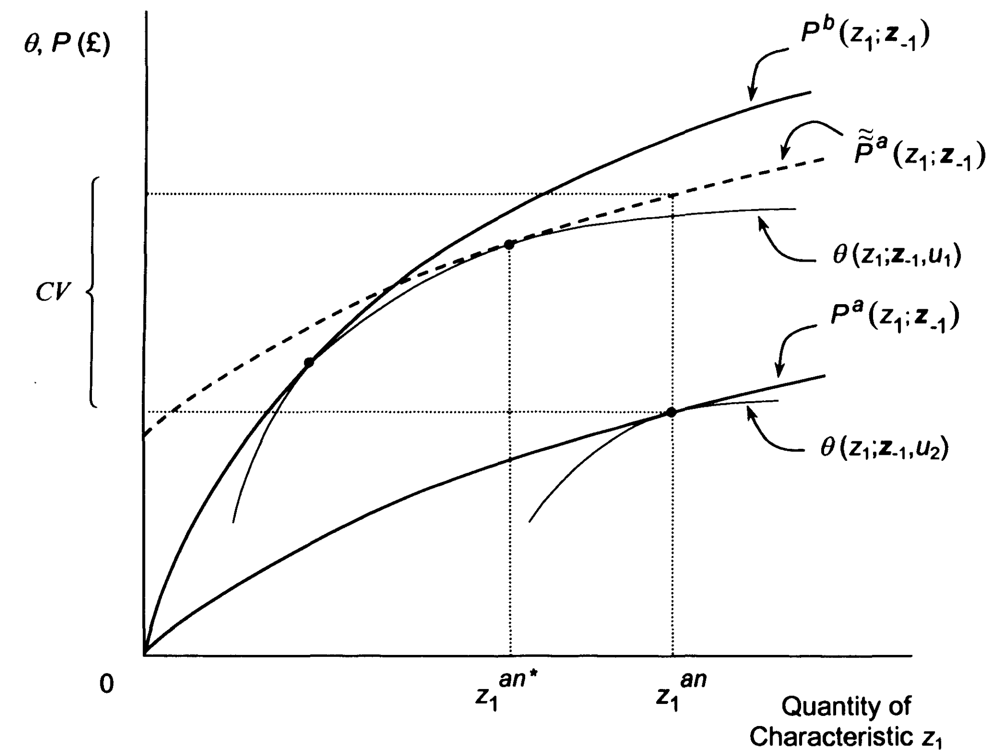
$$W_H = \sum_{h=1}^{H} CS_H = \sum_{h=1}^{H} \begin{pmatrix} \theta\left(z_{1h}^{an}, z_{-1h}^{an}; u_{1h}\right) - \theta\left(z_{1h}^{bo}, z_{-1h}^{bo}; u_{1h}\right) \\ -\left[P^a\left(z_{1h}^{an}, z_{-1h}^{an}\right) - P^b\left(z_{1h}^{bo}, z_{-1h}^{bo}\right)\right] \end{pmatrix} \tag{5}$$

Notice that in comparison with the *QCS* measure, evaluating the *CS* measure of welfare change imposes far greater informational requirements on the researcher. Not only must the researcher be able to evaluate the bid function, but also predict how the HPF will adjust in response to the environmental improvement. Further, the researcher must anticipate the characteristics of the property that each household will choose to rent in response to the new HPF. If the welfare evaluation is to be carried out prior to the environmental improvement, as would be the case in a cost-benefit analysis, these requirements are so onerous as to make the measure practically impossible to evaluate in the real world.

It transpires that even Bartik's *CS* measure of household welfare change is not the most comprehensive measure. In paying out the amount *CS* the household is experiencing a change in income. As their income changes, their optimal choice of residential location will also change. However, in measuring *CS* we have constrained the household to remain in the same residential location. If we relax this constraint then the household can respond optimally by changing their location in response to a change in income. Indeed, allowing the household to respond optimally means that they would be able to pay out a greater amount to achieve the improvement in environmental quality[16].

In Figure 4 we have again illustrated the change in income that would result from paying out a compensating measure as a vertical shift in the HPF. The maximum amount the household would be willing to pay to ensure the change in environmental quality will be the amount that shifts the HPF to the point where it is just tangent with the original bid curve. The point of this tangency would determine the characteristics of the property that the household would decide to rent if it were forced to pay out its maximum WTP to achieve the improvements in environmental standards. We denote the characteristics of this property z^{an*}.

Figure 4: The Compensating Variation measure of the welfare change resulting from an improvement in environmental quality



16 As Palmquist (1986) points out, whenever, we release a constraint on household behaviour we increase their ability to react optimally, thus increasing the quantity of money they would be willing to pay to secure an improvement in environmental quality.

As illustrated in Figure 4, the vertical distance between the hedonic function $P^a(z)$ and the hedonic function as it would appear to the household once it had paid out its maximum WTP, $\tilde{\tilde{P}}^a(z)$, gives a third measure of welfare change that we shall identify as the *compensating variation* (*CV*). This is the measure presented in Palmquist (1986).

CV is the most comprehensive measure of welfare change since it allows the household to react optimally in adjusting to changes in the prices it faces in the market and in adjusting to changes in its own income. The *CV* measure of a welfare change resulting from an improvement in environmental quality will always be greater than the *CS*. However, the informational requirements of the *CV* measure are even greater than those of the *CS* measure. As a consequence we do not consider this measure further.

The various measures of household welfare discussed in this section are summarised in Table 1.

Table 1: Measures of household welfare change

Welfare Measure	Description	Computation of Total Welfare Change for Households	Informational Requirements
Localised:			
No Moving Costs	• No shift in hedonic • Households incur no transaction costs in moving property	$W_H = 0$	• None
Moving Costs	• No shift in hedonic • Households incur transaction costs in moving property	$\sum_{h \in H_1} - tc_h \leq W_H \leq 0$	• Only affected households • Increase in equivalent per period transaction costs
Non-Localised:			
Quantity Compensating Surplus	• Hedonic shifts • Landlords do not change rental on properties • Households remain in their original properties	$W_H = \sum_{h \in H_1} \theta\left(z_{1h}^{ao}, z_{-1h}^{bo}; u_{1h}\right) - \theta\left(z_{1h}^{bo}, z_{-1h}^{bo}; u_{1h}\right)$	• Only affected households • Environmental quality at each affected property before and after improvement • Household bid function
Compensating Surplus	• Hedonic shifts • Landlords change property rents in accordance with the new hedonic • Households relocate to optimal residential locations	$W_H = \sum_{h=1}^{H} \left(\begin{array}{l} \theta\left(z_{1h}^{an}, z_{-1h}^{an}; u_{1h}\right) - \theta\left(z_{1h}^{bo}, z_{-1h}^{bo}; u_{1h}\right) \\ - \left[P^a\left(z_{1h}^{an}, z_{-1h}^{an}\right) - P^b\left(z_{1h}^{bo}, z_{-1h}^{bo}\right)\right] \end{array} \right)$	• All households • Hedonic before and after change • Environmental quality at each affected property before and after improvement • Households choice of residential location in response to new hedonic • Household bid function

5. Changes in Economic Welfare for Landlords

So far we have considered only the demand side of the market. A comprehensive measure of the welfare change resulting from an exogenous environmental improvement should also take account of changes in the profits realised by landlords.

As Bartik (1988) points out, there are four reasons why we would expect a landlord's profits to change after a change in environmental quality;

- If environmental quality at the property's location changes, the property's rental value will change even if the overall hedonic price schedule does not shift.
- Environmental quality changes may affect a landlord's costs (e.g. an increase in air pollution may necessitate more frequent cleaning of the property).
- Any shift in the hedonic function resulting from an environmental improvement affects rents received by all landlords, even those whose property does not directly benefit from the improvement.
- Landlords may respond to all these changes by altering the levels of attributes associated with their property. In so doing they will alter the rental price of the property and also the cost of supplying this property to the market.

As with the discussion for households, we shall work from less comprehensive measures of landlords' welfare change through to a fully comprehensive measure.

5.i. Welfare changes from a localised environmental improvement

To begin with let us consider the welfare impact of a localised environmental improvement. As before, such a change is insufficient to provoke a change in the HPF. This then represents our first assumption.

- **Assumption 1**: The *environmental improvement is localised* and hence does not change the market clearing HPF.
- **Assumption 2**: The landlord cannot independently influence the property's environmental quality. It is entirely determined by exogenous factors.

Assumption 2 results in the corner solution discussed in relation to the right hand panel of Figure 7 in Chapter 1. A similar diagram is reproduced here as Figure 5 where z_1 represents levels of environmental quality. Since the landlord is unable to alter the level of environmental quality through their own actions, the offer curves in Figure 5 reduce to points above the exogenously determined level.

Let us focus on one landlord's property located in the area experiencing the environmental improvement. Initially, this property enjoys a level of environmental quality z_1^b, where, once again the *b* superscript indicates that this is *before* the environmental improvement. Since this is supplied without cost to the landlord, the quantity $\underline{z}_1^b = z_1^b$ is the baseline level of environmental quality. This quantity enters the cost and thence offer functions as an element in the vector $\underline{z}$.

Given the HPF $P^b(z)$, the best the landlord can do is move to the point labelled *X*, coinciding with the offer curve $\phi(z;\underline{z}_1^b,\underline{z}_{-1},\pi^b)$. Here the landlord supplies his property with z_1^b of the environmental attribute and levels of the other property attributes given by the vector, z_{-1}^b. As a result, the landlord can charge a rent of P^b and earns a profit of π^b.

Now, imagine a public programme that increases the level of environmental quality enjoyed at the landlord's property to z_1^a, where the *a* superscript indicates that this is *after* the environmental improvement. Let us make a further assumption;

- **Assumption 3**: The *level of environmental quality does not affect the optimal level of provision of other property characteristics*. Technically, this amounts to assuming that the attribute z_1 does not interact with other arguments in the HPF.

Thus after the environmental improvement, the landlord will maintain the levels of other environmental attributes at z_{-1}^b.

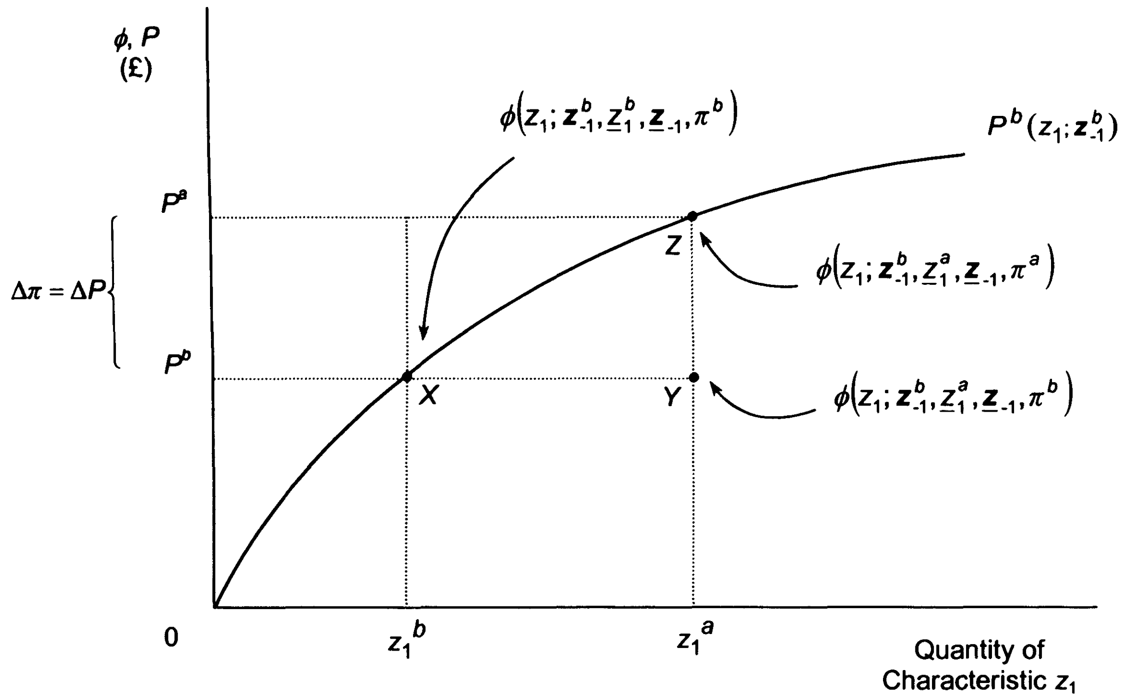
The first welfare measure we consider requires one further assumption;

- **Assumption 4**: The *level of environmental quality does not affect the costs of supplying other property attributes*. Technically this amounts to assuming that the attribute z_1 does not interact with other arguments in the cost function.

Given our four assumptions, measuring the benefits to landlords of the environmental improvement is a relatively straightforward task.

To illustrate the welfare change experienced by a landlord owning a property in the improved area, observe Figure 5.

Figure 5: Landlord welfare change for a localised change in an exogenously determined environmental attribute when costs do not change



Following the environmental improvement, the landlord could continue to charge a rental price of P^b. This would correspond to the point marked Y in Figure 5. There are a number of things to note about this point.

- First, since the improvement is determined by exogenous factors (Assumption 2), the landlord incurs no added cost in supplying the extra environmental quality.
- Second, following the environmental improvement, the landlord continues to supply the other housing attributes at levels given by the vector z^b_{-1} (Assumption 3).
- Finally, the landlord's costs of supplying this vector of other property attributes, z^b_{-1}, will also remain unchanged (Assumption 4).

Since the landlord incurs the same costs after the improvement as before, the profit associated with point *Y* is identical to that associated with point *X*, namely π^b.

Of course, the property now boasts a higher level of environmental quality. Indeed, the landlord is in a position whereby he can increase profits by increasing the rental price of the property. Indeed, given the HPF, the landlord could increase the rental price up to the point marked *Z*. Notice that this increase in rental price adds directly to the landlord's profits. At *Z*, the landlord charges a rental price of P^a and realises a profit π^a.

The welfare measure we seek, therefore, is the difference between profits before the improvement, π^b, and profits after the improvement, π^a. We know from the previous chapter that, provided all else stays the same, the vertical distance between two offer curves equates to the difference in profits associated with the two curves (see Figure 5). Accordingly, the vertical distance *YZ* measures the increase in profits enjoyed by the landlord. Conveniently, this vertical distance is also the difference between the HPF evaluated at the original and improved levels of attribute z_1.

Given our four original assumptions, therefore, the change in profits for the landlord can be written;

$$\Delta\pi = \pi^a - \pi^b = P^b\left(z_1^a, \mathbf{z}_{-1}^b\right) - P^b\left(z_1^b, \mathbf{z}_{-1}^b\right) \tag{6}$$

Of course, we could also derive this result analytically. We know from Equation (17) of Chapter 1 that the profit realised by the landlord for a property with characteristics $\mathbf{z}$ will equal the rental price of such a property minus the cost of providing the property. Thus we could just as easily write;

$$\Delta\pi = \pi^a - \pi^b = \begin{pmatrix} P^b\left(z_1^a, \mathbf{z}_{-1}^b\right) - c\left(z_1^a, \mathbf{z}_{-1}^b; \underline{z}_1^a, \underline{\mathbf{z}}_{-1}\right) \\ -\left[P^b\left(z_1^b, \mathbf{z}_{-1}^b\right) - c\left(z_1^b, \mathbf{z}_{-1}^b; \underline{z}_1^b, \underline{\mathbf{z}}_{-1}\right)\right] \end{pmatrix} \tag{7}$$

Now, we have already assumed that attribute z_1 is provided without cost to the landlord (Assumption 2) and that the level of this attribute has no effect on the costs of providing other property attributes (Assumption 4). As a result, we can conclude that $c\left(z_1^a, z_{-1}^b; \underline{z}_1^a, \underline{z}_{-1}\right)$ and $c\left(z_1^b, z_{-1}^b; \underline{z}_1^b, \underline{z}_{-1}\right)$ take on the same value and fall out of Equation (7) leaving the desired result, Equation (6).

This is, of course, very intuitive. If the improvement allows the landlord to increase the rental price from P^b to P^a but leaves all costs unchanged, the increase in profits for the landlord will simply be the increase in rental price charged on the property.

Given our assumptions, the total welfare gain to landlords will be given by summing Equation (6) across all landlords. Of course, one of those assumptions is that there are no adjustments in the hedonic property market (Assumption 1). Consequently the welfare change will only be experienced by landlords owning properties in the affected area. In the previous section, we denoted this set of properties H_1. Thus, indexing landlords in the market by $l = 1 \text{ to } L$, the welfare change experienced by landlords can be expressed;

$$W_L = \sum_{l \in H_1} P^b\left(z_{1l}^a, z_{-1l}^b\right) - P^b\left(z_{1l}^b, z_{-1l}^b\right) \tag{8}$$

where W_L is the total welfare change experienced by landlords in the market.

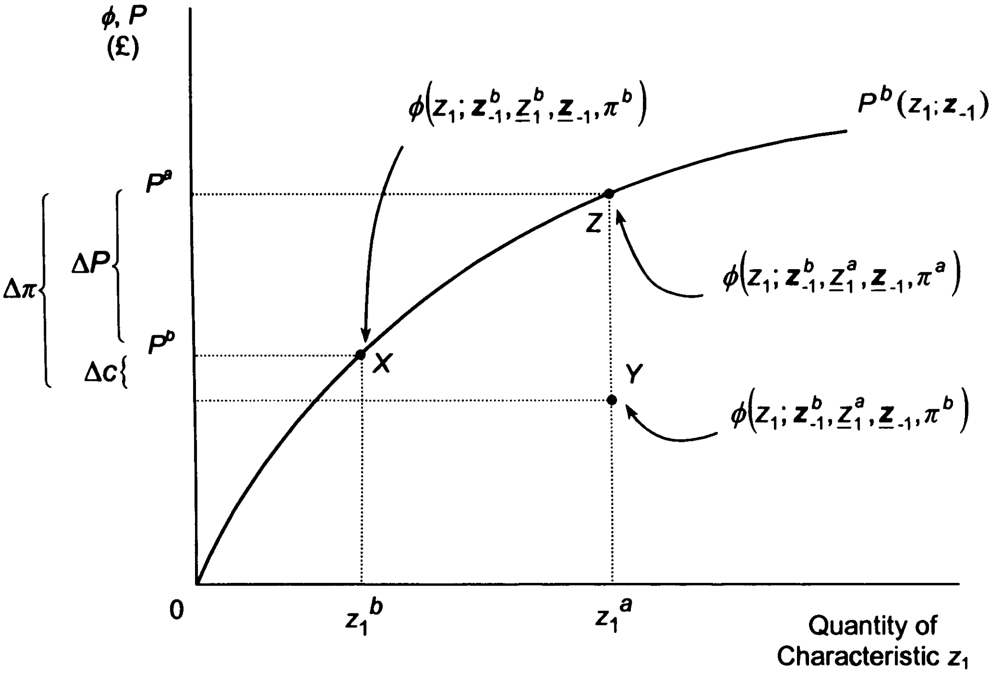
One of the advantages of this welfare measure is that it requires relatively little information. To use this measure, a researcher would simply need an estimate of the HPF and details of the level of the environmental attribute at affected properties before and after the improvement.

Of course, the assumptions made in deriving Equation (8) are very restrictive. For example, consider the situation where Assumption 4 is relaxed such that the level of the environmental attribute z_1 influences the landlords' costs.[17] This case is depicted in Figure 6. Again the environmental improvement has only a local impact (Assumption 1), the level of the attribute is entirely determined by exogenous factors (Assumption 2) and the landlord persists in supplying other property attributes at the same level after the improvement (Assumption 3).

---

[17] For example, reducing air pollution might reduce cleaning and/or repainting costs. Similarly reducing crime might reduce the costs of repairs resulting from vandalism.

Figure 6: Landlord welfare change for a localised change in an exogenously determined environmental attribute when costs change



Before the improvement, the landlord chooses to locate at point *X*. Here the landlord supplies a property with the exogenously determined level of environmental quality z_1^b and chosen levels of other property attributes given by the vector z_{-1}^b. At this combination of attributes the landlord maximises profits by charging a rent P^b of which π^b is profit.

Following an environmental improvement, the level of z_1 is increased to z_1^a at no cost to the landlord. Further according to Assumption 3, the landlord continues to provide other property attributes at the same levels, that is z_{-1}^b. However, by relaxing Assumption 4, we allow for the possibility that the environmental improvement may reduce the cost of providing the other housing attributes at these levels.

Indeed, following the environmental improvement the landlord could locate at point *Y*. Here, the landlord could charge a lower price yet, as a result of cost savings, achieve the same level of profits as previous to the environmental improvement. The vertical distance between *X* and *Y* measures the cost savings brought about by the environmental improvement.

Of course the landlord will not locate at Y. Instead, he will maximise his profits by locating at point Z. Here the landlord charges a rent P^a of which π^a is profit.

The environmental improvement increases the landlord's profits from π^b to π^a. Again, this increase can be measured as the vertical distance between the offer curves, YZ. Notice that allowing for cost changes expands our measure of the welfare gains for landlords. Not only does the landlord enjoy an increase in rent, ΔP, but also experiences a reduction in costs Δc.

Accordingly, this broader welfare measure can be calculated as;

$$\Delta\pi = \Delta P + \Delta c = \left(\begin{array}{c} \left[P^b\left(z_1^a, \mathbf{z}_{-1}^b\right) - P^b\left(z_1^b, \mathbf{z}_{-1}^b\right)\right] \\ + \left[c\left(z_1^b, \mathbf{z}_{-1}^b; \underline{z}_1^b, \underline{\mathbf{z}}_{-1}\right) - c\left(z_1^a, \mathbf{z}_{-1}^b; \underline{z}_1^a, \underline{\mathbf{z}}_{-1}\right)\right] \end{array} \right) \tag{9}$$

Since, this measure continues to assume that there are no adjustments in the hedonic property market the welfare change is only experienced by landlords owning properties in the affected area. Using this measure, the total welfare impact of the environmental improvement is given by;

$$W_L = \sum_{l \in H_1} \left(\begin{array}{c} \left[P^b\left(z_{1l}^a, \mathbf{z}_{-1l}^b\right) - P^b\left(z_{1l}^b, \mathbf{z}_{-1l}^b\right)\right] \\ + \left[c\left(z_{1l}^b, \mathbf{z}_{-1l}^b; \underline{z}_{1l}^b, \underline{\mathbf{z}}_{-1l}\right) - c\left(z_{1l}^a, \mathbf{z}_{-1l}^b; \underline{z}_{1l}^a, \underline{\mathbf{z}}_{-1l}\right)\right] \end{array} \right) \tag{10}$$

Notice that this measure of welfare change is informationally more exacting since it demands that the researcher has knowledge of the landlords cost function.

The two welfare measures that we have developed so far, have both assumed that landlords are not able to influence the level of environmental quality of their properties. Whilst this may be true in the short-term, we have already cited counter examples. For instance, a landlord can change a property's exposure to noise pollution by installing double-glazing.

Our next task, therefore, is to relax Assumption 2 and consider the situation where the level of environmental quality is not entirely determined by exogenous factors. For now, however, we maintain Assumption 3. That is, following an environmental improvement, we allow landlords to alter the level of environmental quality of their properties but not alter the levels of other property attributes.

The complex pattern of responses is laid out in Figure 7. In the first instance the landlord is faced by the HPF $P^b(\cdot)$ and the exogenously determined level of the environmental attribute $\underline{z}_1^b$. To illustrate let us assume that z_1 is the level of crime in the area. Faced with these two restrictions, the landlord maximises profits by investing in private goods that expand the level of attribute z_1 to z_1^{bn}. Here the superscript *n* indicates the *new* level of the property attribute once the investments have been undertaken. For instance the landlord could further reduce the risk of crime by installing a burglar alarm monitored by a private security company. Following these investments, the landlord achieves point *W* where the rental value of the property is P^b and the landlord earns a profit of π^b.

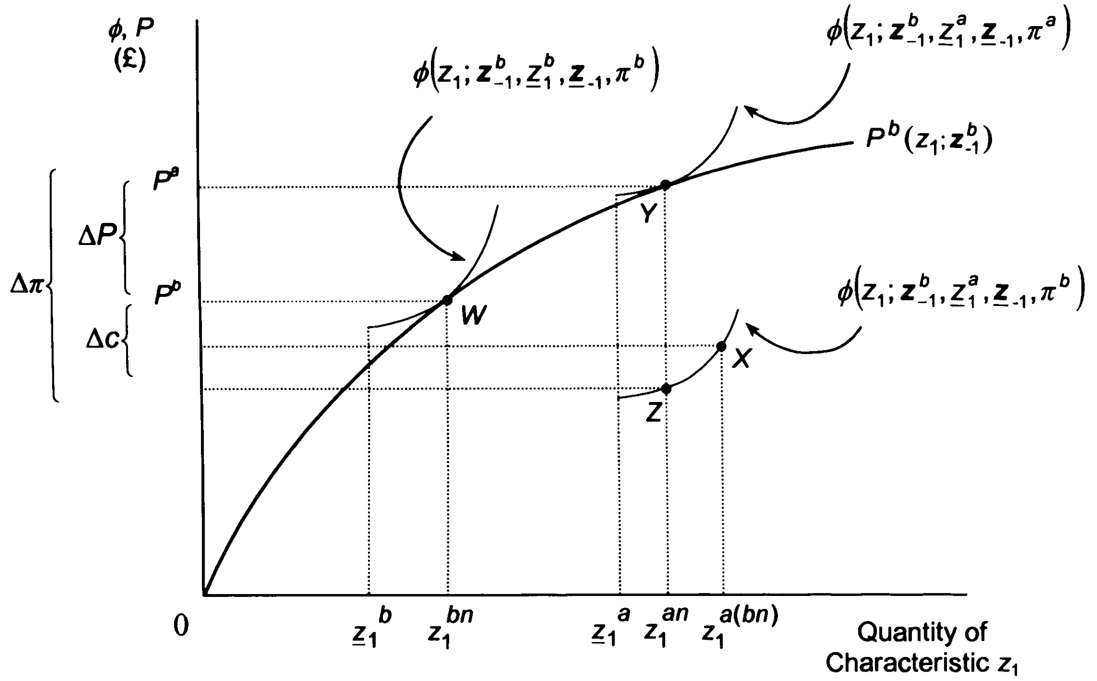
Now let us consider a public programme that leads to an increase in the exogenously supplied level of z_1 from $\underline{z}_1^b$ to $\underline{z}_1^a$. In our example, the level of criminal activity in the area falls. For the sake of argument, imagine that the landlord did not adjust to this change. In our example, the landlord might continue to employ the private security firm despite the fact that crime risks in the area have fallen. Following the change the landlord's property would boast a level of environmental quality given by $z_1^{a(bn)}$ where the superscript $a(bn)$ indicates that this is the level of provision *after* the change but whilst maintaining the *new* level of property investments undertaken *before* the change.

Thus if the landlord wished to maintain the same level of profit as previous to the change, he would end up at point *X* which lies on the new offer curve providing the original level of profit, π^b.

Notice that, as in the previous scenario, the increased environmental quality has resulted in immediate reductions in the costs of providing other housing attributes. Indeed, the vertical distance between *W* and *X* measures the cost savings brought about by the environmental improvement.

Of course *X* is by no means the landlord's optimal location. Indeed, given $P^b(\cdot)$ and the exogenously determined level of the environmental attribute $\underline{z}_1^a$, the landlord would be best advised to increase the rent on the property and consider the potential benefits of changing the property's level of environmental quality.

Figure 7: Landlord welfare change for a localised change in environmental attribute



In Figure 7 the best the landlord could do would be to relocate to point Y. Here, the level of the environmental attribute z_1 has been altered to z_1^{an} and the landlord maximises profits at π^a by charging a rental of P^a. Continuing our example, in response to the fall in crime in the area, the landlord may decide to increase the rent on the property whilst terminating his employment of the private security company.

Once again, the increase in the landlord's profits will be the vertical distance between Y and the point on the equivalent offer curve delivering the original level of profits, point Z. In Figure 7, therefore, the increase in the landlord's profits is the distance ZY.

Again this increase in profits can be decomposed into a change of price and a change in costs according to;

$$\Delta\pi \;=\; \Delta P + \Delta c \;=\; \begin{pmatrix} \left[P^b\left(z_1^{an}, z_{-1}^b\right) - P^b\left(z_1^{bn}, z_{-1}^b\right)\right] \\ + \left[c\left(z_1^{bn}, z_{-1}^b; \underline{z}_1^b, \underline{z}_{-1}\right) - c\left(z_1^{an}, z_{-1}^b; \underline{z}_1^a, \underline{z}_{-1}\right)\right] \end{pmatrix} \tag{10}$$

This measure is broader than those discussed previously since it allows for the adjustments that landlords make in the provision of environmental quality after the exogenous change. Since we are still assuming a localised environmental improvement, this broader measure will still only be defined for properties in the affected area. The total welfare change is given by;

$$W_L = \sum_{l \in H_1} \left(\begin{array}{c} \left[P^b\left(z_{1l}^{an}, z_{-1l}^b\right) - P^b\left(z_{1l}^{bn}, z_{-1l}^b\right) \right] \\ + \left[c\left(z_{1l}^{bn}, z_{-1l}^b; \underline{z}_{1l}^b, \underline{z}_{-1l}\right) - c\left(z_{1l}^{an}, z_{-1l}^b; \underline{z}_{1l}^a, \underline{z}_{-1l}\right) \right] \end{array} \right) \tag{11}$$

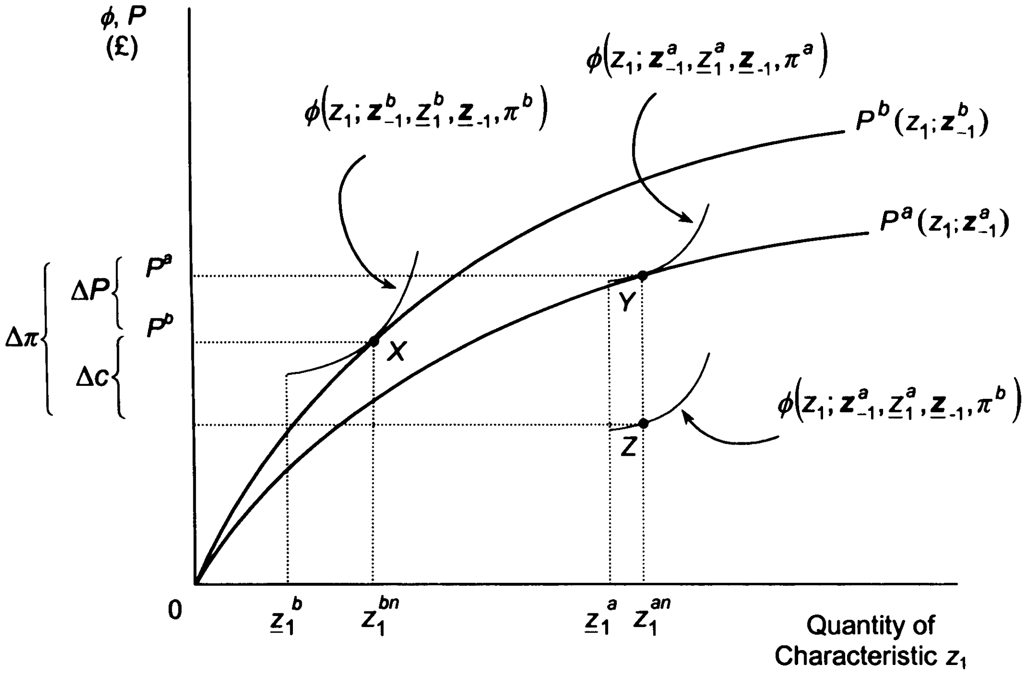
5.ii. *Welfare changes from a non-localised environmental improvement*

The final welfare measure we discuss relaxes all four assumptions simultaneously. Here the environmental improvement is assumed to be substantial enough to result in a shift in the HPF. Further, unlike the measure described by Equation (11), we allow for the fact that the landlord may decide to change the levels of provision of all the housing attributes as a result of the environmental improvement and subsequent shift in the HPF. This case is depicted in Figure 8.

The landlord starts off with an exogenously determined level of environmental quality $\underline{z}_1^b$ and baseline levels of other property attributes given by the vector $\underline{z}_{-1}$. In the first instance the landlord is faced by the HPF $P^b(\cdot)$. In order to maximize profits the landlord wishes to move to point *X* by altering the environmental quality of the property to z_1^{bn} and the levels of other property attributes z_{-1}^b. Here the landlord can charge a price of P^b and earns profits from the property of π^b.

Now a public programme results in an environmental improvement in the urban area. At the landlord's property this manifests itself as an increase in the exogenously determined level of environmental quality from $\underline{z}_1^b$ to $\underline{z}_1^a$. However, this is not a merely localised change. Indeed, the set of prices given by the old HPF would no longer clear the market. In response to the environmental improvement, the market adjusts, establishing equilibrium at the new HPF given by $P^a(\cdot)$.

Figure 8: Landlord welfare change for a non-localised change in an environmental attribute



The landlord is faced by a number of simultaneous changes;

- environmental quality at their property increases
- the environmental improvement reduces the costs of providing different combinations of property attributes
- the HPF shifts

In response the landlord maximises profits by altering the provision of environmental quality to z_1^{an} and the levels of other property attributes to z_{-1}^{a}, point *Y*. Notice that we have allowed for the fact that it may be optimal to adjust the level of all housing attributes in response to the environmental improvement.

Following the same argument as that used previously, the relevant welfare measure is the vertical distance between the points marked *Z* and *Y*.

This measure is the landlords' equivalent to the Compensating Surplus measure defined for households. As with that measure, the landlord is allowed to respond optimally to the change in environmental quality and the shift in the HPF. For this reason we label this comprehensive welfare measure the *Compensating Profit* (*CP*). In mathematical terms it is defined as;

$$CP \;=\; \Delta\pi \;=\; \Delta P + \Delta c \;=\; \left(\begin{array}{c} \left[P^a\left(z_1^{an}, z_{-1}^a\right) - P^b\left(z_1^{bn}, z_{-1}^b\right)\right] \\ + \left[c\left(z_1^{bn}, z_{-1}^b; \underline{z}_1^b, \underline{z}_{-1}\right) - c\left(z_1^{an}, z_{-1}^a; \underline{z}_l^a, \underline{z}_{-1}\right)\right] \end{array} \right) \tag{12}$$

If $\Delta\pi$ is negative then the change in environmental quality reduces the welfare of the landlord. If $\Delta\pi$ is positive then the change in environmental quality increases the welfare of the landlord.

Since all landlords are assumed to respond to the shift in the HPF the total welfare benefits of the environmental improvement will include a measure for each of the H landlords in the urban area;

$$W_L \;=\; \sum_{l=1}^{H} CP_l \;=\; \sum_{l=1}^{H} \left(\begin{array}{c} \left[P^a\left(z_{1l}^{an}, z_{-1l}^a\right) - P^b\left(z_{1l}^{bn}, z_{-1l}^b\right)\right] \\ + \left[c\left(z_{1l}^{bn}, z_{-1l}^b; \underline{z}_{1l}^b, \underline{z}_{-1l}\right) - c\left(z_{1l}^{an}, z_{-1l}^a; \underline{z}_{1l}^a, \underline{z}_{-1l}\right)\right] \end{array} \right) \tag{13}$$

Notice that the informational requirements of the *CP* measure are extremely onerous. Not only must the researcher be able to predict how the HPF will change in response to a non-localised change in environmental quality, but must also be able to predict the optimal response of each landlord to the change in market conditions.

Table 2 summarises the various measures of landlord welfare change described in this section.

Table 2: Measures of landlord welfare change

Welfare Measure	Description	Computation of Total Welfare Change for Landlords	Informational Requirements
Localised:			
Exogenous Attribute, no Cost Changes	• No shift in hedonic • Rent increase for improved properties	$W_L = \sum_{l \in H_1} P^b\left(z_{1l}^a, \mathbf{z}_{-1l}^b\right) - P^b\left(z_{1l}^b, \mathbf{z}_{-1l}^b\right)$	• Only affected landlords • Environmental quality before and after change • Original hedonic
Exogenous Attribute, with Cost Changes	• No shift in hedonic • Landlords in improved areas experience cost changes • Rent increase for improved properties	$W_L = \sum_{l \in H_1} \begin{pmatrix} P^b\left(z_{1l}^a, \mathbf{z}_{-1l}^b\right) - P^b\left(z_{1l}^b, \mathbf{z}_{-1l}^b\right) \\ + \left[c\left(z_{1l}^b, \mathbf{z}_{-1l}^b; \underline{z}_{1l}^b, \underline{\mathbf{z}}_{-1l}\right) - c\left(z_{1l}^a, \mathbf{z}_{-1l}^b; \underline{z}_{1l}^a, \underline{\mathbf{z}}_{-1l}\right)\right] \end{pmatrix}$	As previous, plus: • Changes in exogenous levels of other attributes • Landlord cost function
Any attribute	• No shift in hedonic • Landlords in improved areas experience cost changes • Landlords optimise level of provision of environmental quality attribute • Rent change for improved properties	$W_L = \sum_{l \in H_1} \begin{pmatrix} \left[P^b\left(z_{1l}^{an}, \mathbf{z}_{-1l}^b\right) - P^b\left(z_{1l}^{bn}, \mathbf{z}_{-1l}^b\right)\right] \\ + \left[c\left(z_{1l}^{bn}, \mathbf{z}_{-1l}^b; \underline{z}_{1l}^b, \underline{\mathbf{z}}_{-1l}\right) - c\left(z_{1l}^{an}, \mathbf{z}_{-1l}^b; \underline{z}_{1l}^a, \underline{\mathbf{z}}_{-1l}\right)\right] \end{pmatrix}$	As previous, plus: • Landlords' choices of environmental quality attribute after improvement
Non-Localised:			
Compensating Profit	• Hedonic shifts • Landlords in improved areas experience cost changes • Landlords optimise property attributes • Rent change for all properties	$W_L = \sum_{l=1}^{H} \begin{pmatrix} P^a\left(z_{1l}^{an}, \mathbf{z}_{-1l}^a\right) - P^b\left(z_{1l}^{bn}, \mathbf{z}_{-1l}^b\right) \\ + \left[c\left(z_{1l}^{bn}, \mathbf{z}_{-1l}^b; \underline{z}_{1l}^b, \underline{\mathbf{z}}_{-1l}\right) - c\left(z_{1l}^{an}, \mathbf{z}_{-1l}^a; \underline{z}_{1l}^a, \underline{\mathbf{z}}_{-1l}\right)\right] \end{pmatrix}$	As previous, plus: • All landlords • Landlords' choices of all attributes after improvement • Hedonic before and after change

6. Combining Household and Landlord Welfare Measures

The total benefits to households and landlords resulting from an environmental improvement are found simply by adding W_H to W_L. Of course, this total welfare measure will depend on which assumptions are made and hence which measures are chosen to represent the households' and landlords' welfare changes.

We should note that such welfare estimates measure the benefits to both households and landlords for changes in environmental quality in their residential location. However, they ignore the;

- benefits to visitors that travel by the improved area.
- benefits to those who work in the improved area[18].
- costs of causing the environmental improvement. For example, no account is taken of the costs to industry of reducing emissions or the cost to the tax payer of traffic calming schemes designed to reduce traffic noise.

In the simplest case, the environmental improvement is a localised phenomenon that has no impact on the HPF. If we assume that households incur no moving costs then they will relocate to a property offering the attributes of their original location prior to the improvement and experience no welfare change. Further, if we assume that landlords cannot affect the level of environmental quality at their properties, that the level of environmental quality does not influence the optimal level of provision of other attributes and that their costs of providing other property attributes are unaffected by the improvement, then the welfare gain for landlords is simply the change in the rental price of their properties. The total welfare change is given by the sum of Equation (8) and the upper bound of Equation (1);

$$W_L + W_H \;=\; \sum_{l \in H_1} P^b\left(z_{1l}^a, \mathbf{z}_{-1l}^b\right) - P^b\left(z_{1l}^b, \mathbf{z}_{-1l}^b\right) \;+\; 0 \qquad (13)$$

[18] We could attempt to measure the benefits to these individuals in other hedonic markets such as that for office space or for labour.

Under certain restrictive assumptions, therefore, the total welfare change can be measured as the change in the price of affected properties. What is more, to calculate this measure requires only two pieces of information;

- the current HPF.
- the level of environmental quality at each affected property before and after the environmental improvement.

For any one market, welfare changes as measured by Equation (13) should be relatively simple to estimate. Unfortunately, it is not possible to transfer such estimates to different property markets. Remember from Chapter 1 that HPFs differ across property markets according to the particular conditions of supply and demand. A welfare measure calculated using the HPF in one particular market would only be relevant to that market. It would make no sense to transfer such estimates across markets.

Of course, Equation (13) is by no means a comprehensive measure of the welfare change associated with a localised change in environmental quality. Indeed, by relaxing some of the assumptions underlying Equation (13) we could expand our measure of the welfare gain. For example, we might wish to allow for the fact that households face transaction costs when moving properties, that landlords might wish to optimally adapt the level of environmental quality at their properties and that changes in environmental quality might affect the costs of providing other property attributes. In this case our welfare measure would be the sum Equation (11) and the lower bound of Equation (1);

$$W_L + W_H = \sum_{l \in H_1} \left(\begin{array}{c} \left[P^b\left(z_{1l}^{an}, \mathbf{z}_{-1l}^{b}\right) - P^b\left(z_{1l}^{bn}, \mathbf{z}_{-1l}^{b}\right) \right] \\ + \left[c\left(z_{1l}^{bn}, \mathbf{z}_{-1l}^{b}; \underline{z}_{1l}^{b}, \underline{\mathbf{z}}_{-1l}\right) - c\left(z_{1l}^{an}, \mathbf{z}_{-1l}^{b}; \underline{z}_{1l}^{a}, \underline{\mathbf{z}}_{-1l}\right) \right] \end{array} \right) - \sum_{h \in H_1} tc_h \qquad (14)$$

Whilst this may be a more comprehensive measure of the welfare change it is also considerably harder to estimate. Compared to Equation (13) the researcher would need to estimate the moving costs for each household affected by the environmental change, the landlords' cost function and the adaptations made by landlords to the environmental quality attribute following the improvement. Indeed, attempting to estimate Equation (13) prior to a change in environmental quality is almost an impossible task.

In the extreme, we could relax all assumptions and allow for changes in environmental quality that are non-localised and precipitate alterations in the HPF. Ignoring transaction costs, this measure would be derived by adding Equation (13) to Equation (5);

$$
\begin{aligned}
TSB &= \sum_{l} CP_l + \sum_{h} CS_h \\
&= \sum_{l=1}^{H} \begin{pmatrix} \left[P^a\left(z_{1l}^{an}, z_{-1l}^{a}\right) - P^b\left(z_{1l}^{bn}, z_{-1l}^{b}\right)\right] \\ + \left[c\left(z_{1l}^{bn}, z_{-1l}^{b}; \underline{z}_{1l}^{b}, \underline{z}_{-1l}\right) - c\left(z_{1l}^{an}, z_{-1l}^{a}; \underline{z}_{1l}^{a}, \underline{z}_{-1l}\right)\right] \end{pmatrix} \\
&\quad + \sum_{h=1}^{H} \begin{pmatrix} \theta\left(z_{1h}^{an}, z_{-1h}^{an}; u_{1h}\right) - \theta\left(z_{1h}^{bo}, z_{-1h}^{bo}; u_{1h}\right) \\ - \left[P^a\left(z_{1h}^{an}, z_{-1h}^{an}\right) - P^b\left(z_{1h}^{bo}, z_{-1h}^{bo}\right)\right] \end{pmatrix}
\end{aligned}
\tag{15}
$$

This would give us a comprehensive measure of the welfare change and hence is labelled the *Total Social Benefits* (*TSB*) of the change in environmental quality. Unfortunately, *TSB* would be almost impossible to calculate. To assess Equation (15) researchers would require detailed knowledge of how the HPF would be affected by changes in environmental quality and how every household and landlord would respond to changes in environmental quality and changes in the hedonic price schedule. As such, Equation (15) is of little use to practitioners attempting to measure the benefits derived from a program designed to change environmental quality in an urban area.

7. A Quantifiable Lower Bound

Since the informational requirements for measuring *TSB* are prohibitive, economists have looked to define a simpler measure that might lend itself to estimation in the real world. It turns out that one such measure is the sum of *QCS* measures presented in Equation 3. All that is required to calculate this measure is knowledge of the bid function of households in the affected area, details of their current residential choices and information on the level of environmental change experienced by each household.

Encouragingly, Bartik (1988) has given a theoretical justification for choosing to measure the welfare changes resulting from a change in environmental quality as the sum of households' *QCS*. He shows that the sum of *QCS* across all affected households provides a lower bound estimate of the *TSB*.

Bartik's intuitive proof involves partitioning the welfare changes affecting households and landlords into a series of three stages. Whilst these stages help in the analysis of welfare changes they are not meant to represent a realistic sequence of events. The three stage decomposition is presented in Table 3.

In the *first stage*, some or all of the residential locations in the urban area experience an improvement in environmental quality. It is assumed that neither landlords, nor households nor the hedonic market adjust in response to this change.

- Since households do not move property, the benefit to households will be simply their WTP for the environmental improvement at their original location. This is the *QCS* measure presented in Figure 2 and Equation (2).
- Since landlords do not change rents or adjust the attributes of their properties, they will only be affected by the change in environmental quality if it affects their costs. Since we assume they make no changes to their properties at this stage, the measure of cost savings is that given by the vertical distance between *W* and *X* in Figure 7.

In the *second stage*, the HPF shifts precipitating a change in the rental price for each property. However, at this stage households and landlords are constrained to their original location and supply choices. As such the change in rents simply acts so as to transfer money from one to the other. Indeed, whatever the pattern of rent changes in the second stage, there is no overall welfare effect.

Notice, however, that whilst there is no change in the aggregate welfare change in stage two, welfare changes for each individual household and landlord may be positive or negative depending on the particular pattern of rent changes.

In the *third stage*, households and landlords respond to the new HPF. Households will move to the property that offers them the highest possible utility. This must be at least as beneficial as remaining in the original property since they could always opt not to move house. Similarly, landlords may adjust the attributes of their properties.

Clearly, any such adjustments must increase profits since the landlord could just as well choose to leave the property as it is. Hence, in stage 3, both households and landlords must enjoy an increase in welfare.

This is not to say that every household and landlord experiences an increase in welfare over all three stages. Whilst households and landlords only benefit in stages 1 and 3, they may just as well lose as gain in the rental changes isolated in stage 2.

As shown in Table 3, summing all three stages for households results in the total welfare gains given by the sum of household *CS's* given in Equation (5). Similarly, summing all three stages results in the sum of landlords *CP*'s given in Equation (13). Thus the three stage decomposition, whilst not reflecting the simultaneous nature of responses to the change in environmental quality, accurately represents the overall change in welfare.

The insight of Bartik's decomposition is to isolate all individual welfare losses as price changes in stage 2. Since price changes simply represent pecuniary transfers between agents in the property market, these losses must be offset by equivalent gains elsewhere. In other words, when we are interested in the aggregate welfare change, we can ignore the losses incurred by certain landlords and households by netting these out as a price change.

As a result *TSB*, that is the total welfare change experienced by all households and landlords in the urban area, can be regarded as the sum of the four non-negative values defined in stages one and three. In words, these are;

1. WTP of households at improved locations to enjoy the change in environmental quality whilst staying in their original property ($\sum_{h \in H_1} QCS_h$)
2. Cost savings for landlords at stage 1
3. Household utility gains from relocation at stage 3
4. Landlord profit gains from changes in supply at stage 3

Since all four values are non-negative, $\sum_{h \in H_1} QCS_h$ must also be a lower bound to *TSB*.

Table 3: A decomposition of the welfare effects of a change in environmental quality from Bartik (1988)

	Benefits at Various Stages		
	Households	Landlords	Net Efficiency Benefits
Stage 1: Amenity changes, no adjustment or rent change	$\sum_h \left[\theta\left(z_{1h}^{ao}, z_{-1h}^{bo}; u_{1h}\right) - \theta\left(z_{1h}^{bo}, z_{-1h}^{bo}; u_{1h}\right)\right]$ Household WTP at original location: zero for unimproved sites, positive for improved sites	$\sum_l -\left[c\left(z_{1l}^{a(bn)}, z_{-1l}^{b}; \underline{z}_{1l}^{a}, \underline{z}_{-1l}\right) - c\left(z_{1l}^{bn}, z_{-1l}^{b}; \underline{z}_{1l}^{b}, \underline{z}_{-1l}\right)\right]$ Landlord cost savings: assumed non-negative for improved sites, zero for unimproved sites	Sum of all households' WTP plus all landlords' cost savings
Stage 2: Rent Change	$\sum_h -\left[P^a\left(z_{1h}^{ao}, z_{-1h}^{bo}\right) - P^b\left(z_{1h}^{bo}, z_{-1h}^{bo}\right)\right]$ Rent change at both improved and unimproved sites	$\sum_l P^a\left(z_{1l}^{a(bn)}, z_{-1l}^{b}\right) - P^b\left(z_{1l}^{bn}, z_{-1l}^{b}\right)$ Rent change at both improved and unimproved sites	Zero efficiency benefits; pecuniary transfer between households and landlords
Stage 3: Adjustment	$\sum_h \theta\left(z_{1h}^{an}, z_{-1h}^{an}; u_{1h}\right) - P^a\left(z_{1h}^{an}, z_{-1h}^{an}\right) - \left[\theta\left(z_{1h}^{ao}, z_{-1h}^{bo}; u_{1h}\right) - P^a\left(z_{1i}^{ao}, z_{-1h}^{bo}\right)\right]$ Measure of household utility increase from adjustment, for households originally at both improved and unimproved sites	$\sum_l P^a\left(z_{1l}^{an}, z_{-1l}^{a}\right) - c\left(z_{1l}^{an}, z_{-1l}^{a}; \underline{z}_{1l}^{a}, \underline{z}_{-1l}\right) - \left[P^a\left(z_{1l}^{a(bn)}, z_{-1l}^{b}\right) - c\left(z_{1l}^{a(bn)}, z_{-1l}^{b}; \underline{z}_{1l}^{a}, \underline{z}_{-1l}\right)\right]$ Landlord profit increase from adjustment to new hedonic: applies to landlords at all sites	Net gain from adjustment must be non-negative for all
Sum of three stages	$\sum_h \theta\left(z_{1h}^{an}, z_{-1h}^{an}; u_{1h}\right) - \theta\left(z_{1h}^{bo}, z_{-1h}^{bo}; u_{1h}\right) - \left[P^a\left(z_{1h}^{an}, z_{-1h}^{an}\right) - P^b\left(z_{1h}^{bo}, z_{-1h}^{bo}\right)\right]$ Net household gain: sum over all households, Equation (5) in text	$\sum_h P^a\left(z_{1l}^{an}, z_{-1l}^{a}\right) - c\left(z_{1l}^{an}, z_{-1l}^{a}; \underline{z}_{1l}^{a}, \underline{z}_{-1l}\right) - \left[P^b\left(z_{1l}^{bn}, z_{-1l}^{b}\right) - c\left(z_{1l}^{bn}, z_{-1l}^{b}; \underline{z}_{1l}^{b}, \underline{z}_{-1l}\right)\right]$ Net landlord gain, sum over all landlords, Equation (13) in text	Sum of 1st and 2nd columns is same as Equation (15)

This is an extremely important insight since it gives us a good theoretical reason for using $\sum_{h \in H_1} QCS_h$ to measure the welfare change resulting from an environmental improvement. There are a number of reasons why this might be desirable.

- Since the *QCS* measure does not require information on how the market price or agents in the market adjust to a change in market conditions, it can be calculated in advance of a public programme to improve environmental quality.
- *QCS* is a measure of household welfare change. Consequently using the sum of *QCS*s as a lower bound estimate of *TSB* removes the need to examine the supply side of the market. Researchers can ignore the considerable difficulties associated with estimating landlord cost and offer functions.
- *QCS* is only defined for households in an affected area. As a result, the researcher only requires information on which households will be affected by the environmental improvement and the extent of improvement enjoyed by each.
- *QCS* is based solely on underlying preferences for environmental quality as captured in the bid function. The measure is not particular to a specific property market. Indeed, if a researcher could derive the bid function from one market then this could be used to evaluate the *QCS* in another property market, provided the researcher was prepared to assume that preferences for environmental quality were stable across the two markets.

Clearly, using the sum of households' *QCS* as a lower bound approximation to the *TSB* makes it practical to carry out *ex-ante* assessments of the welfare gains from environmental improvements. The accuracy of this approximation will depend on the size of the values taken by the other three elements of *TSB* isolated in Bartik's analysis. Certainly, the approximation will tend to be more accurate when the environmental change is less extensive as the benefits of household relocation and landlord change in supply will tend to be smaller.

8. Conclusions

This chapter has demonstrated how the benefits of an environmental improvement can be measured in the property market. In the simplest case, the environmental improvement is a localised phenomena that causes no change in the HPF. If households can move freely and landlords do not enjoy cost savings and are constrained not to alter the supply of property attributes, then the welfare benefits of the improvement accrues to landlords as the *change in the rental price* of their properties (Equation 13).

This measure is easy to calculate for any property market for which the HPF is known. Unfortunately, the fact that the measure is based on the unique HPF of a particular market means that there is no theoretical substance to transferring such values across property markets.

Clearly, estimating the welfare change of an environmental improvement by the increase in prices of affected properties is to impose severe restrictions on the reactions of the economic agents in the market to the improvement. Indeed, a completely comprehensive measure of the welfare benefits of an environmental improvement is given by the *TSB* measure (Equation 15).

However, the *TSB* measure is little more than a theoretical construct. To estimate such a measure researchers would require detailed knowledge of how the equilibrium HPF would be affected by changes in environmental quality and how households' and landlords' choices would respond to both changes in environmental quality and changes in the hedonic price schedule.

Unfortunately, hedonic market equilibria are too complex to derive satisfactory analytical solutions by which to predict such outcomes. Indeed, the *TSB* measure is almost impossible to calculate *ex-ante*, making it of little use to practitioners attempting to measure the potential benefits of a program seeking to change environmental quality in an urban area.

Nevertheless, in an important analysis, Bartik (1988) showed how a third measure the *QCS,* when summed over all households directly affected by the change in environmental quality, could always be taken as a lower bound to the *TSB*. There are a number of reasons why using the *QCS* measure might be desirable. In particular, the *QCS* measure is based solely on the household bid function. As a result, it is not necessary to consider the supply side of the market nor predict market conditions following environmental change. Further, the *QCS* measure is not particular to a specific property market. Indeed, if a researcher could derive the bid function from one market then, provided the researcher was prepared to assume that preferences for environmental quality were stable across the two markets, this could be used to evaluate the *QCS* in another property market.

In Part 3 of this thesis, therefore, we investigate the possibilities for deriving estimates of the bid function from which the *QCS* measure of welfare change can be derive.

PART 2

EMPIRICAL ESTIMATION OF THE HEDONIC PRICE FUNCTION

CHAPTER 3: THE HEDONIC PRICE FUNCTION IN THEORY AND PRACTICE

1. Introduction

The five chapters that make up Part 2 of this thesis concern themselves with the estimation of the *hedonic price function* (HPF);

$$P = P(\mathbf{z}) \qquad (1)$$

the function that relates a property's characteristics, represented by the vector of attribute levels $\mathbf{z}$, to the price at which that property sells in the market, P.

Chapter 4 provides a description of the data set for which the HPF is to be estimated. The data relate to residential house sales in the City of Birmingham in the UK. As described in that chapter, the data set is remarkably comprehensive compiling information from numerous sources with the aid of geographical information systems (GIS). Chapters 5, 6 and 7 concern themselves with the actual estimation of the HPF for the City of Birmingham data set. The particular focus of the analysis is the identification of the implicit price of transport related noise; that is the amount by which property prices decline for each extra decibel of noise to which they are exposed.

One key research theme introduced in Chapter 4 and further developed in Chapter 6 is the application of data-driven techniques for partitioning property market data into relatively homogeneous subsets. For example, each subset might identify groups of properties exhibiting similar structural characteristics. Alternatively each subset might identify properties inhabited by residents with similar socioeconomic attributes. Whilst numerous previous studies have sort to partition property market data, the innovation presented in this thesis is the application of techniques of *model-based cluster analysis*. The great advantage of these techniques is that they provide an independent statistical indication of the nature and number of homogenous subsets to be found in the data. These techniques represent the cutting edge in cluster analysis and have not previously been employed in hedonic analysis nor, as far as can be ascertained by the author, in any other field of empirical economics. Whilst Chapter 4 presents a standard application of model-based clustering, Chapter 6 adds another level of sophistication, allowing for outliers in the data; that is properties that cannot easily be allocated to a particular partition.

Following the substantial literature in this field, Chapter 4 motivates the partitioning of the data through the assumed presence of market segments. In particular, it is argued that a significant difference in the HPFs of different partitions provides evidence that these partitions identify different market segments. Chapter 6 readdresses this motivation and represents a development in thinking over that presented in Chapter 4. In this later Chapter it is argued that the conditions that lead to market segmentation are unlikely to hold in the property market for one urban area. Rather it is argued that significant differences in the HPF between partitions of the data are evidence of substantial non-linearity in the HPF for Birmingham. Indeed, this observation provides an alternative and original justification for partitioning property market data. In this case, the data are partitioned to facilitate an estimation strategy that seeks to locally approximate a possibly highly non-linear equilibrium HPF.

A second major concern in Part 2 of this thesis is the econometric estimation of the HPF. Two major themes are developed through these chapters. The first theme concerns the econometric specification of the HPF. The second concerns spatial autocorrelation in the regression residuals.

In Chapter 5 sophisticated econometric techniques are used to analyse the Birmingham data. In particular, a well-known semiparametric estimator (Robinson, 1988) is used to introduce flexibility into the specification of the HPF. Improvements on previous applications of this model are made by allowing both selected property characteristics and the influence of location to enter the econometric model nonparametrically. Analysis of the regression residuals from the hedonic price regressions reveals evidence of spatial correlation. As such, a general method of moments estimator proposed by Kelejian and Prucha (1999) is used in a second stage regression. This estimator specifically accounts for spatial autocorrelation in regression residuals returning robust estimates of the model's parameters and their variance. As far as the author is aware, this is the first application to combine semiparametric methods with the Kelejian and Prucha GMM estimator to provide robust estimates of the parameters of a HPF.

Chapter 7 concerns itself with issues of spatial autocorrelation of regression residuals. In particular, it is argued that spatial autocorrelation of residuals is evidence of omitted variables describing spatial features influencing property prices that are not observed by the researcher. As such an estimation approach championed by Gibbons and Machin (2001) is adopted. This approach accounts nonparametrically for omitted spatial variates by spatially smoothing the data. A major innovation of this Chapter is the introduction of a procedure (similar to that proposed by Ellner and Seifu, 2002, in a more general context) that uses the data to prescribe the optimal areal extent of smoothing.

This Chapter, Chapter 3, brings together many of the theoretical and econometric issues relevant to the estimation of Equation (1) and held in common by all the Chapters in Part 2 of this thesis. Further, since the key motivation of Chapter 5 is to identify the impact of noise pollution on property prices, it provides a brief review of other hedonic analyses that have dealt with this issue. These results provide a point of comparison against which the results of the research reported here can be evaluated.

2. Equilibrium in Real World Property Markets

Chapter 1 described the theoretical model of the property market that forms the foundation of empirical hedonic analyses. In this Section we explore the assumptions that underpin the model and assess the extent to which it is appropriate to derive a hedonic price schedule from data on house prices in real property markets.

The fundamental assumption of the theoretical models is that the property market attains equilibrium. That is, the model assumes that the market settles on a set of prices at which households cannot increase their utility by moving to another property (given the constraints of their budgets) and landlords are unable to increase their profits by changing the characteristics of their properties. If we can assume that the property market settles on such an equilibrium then house prices from a cross-sectional dataset should provide the information required to estimate the HPF.

As Freeman (1993) discusses, however, there are a number of reasons why the market may fail to be in *constant* equilibrium;

1. *Households do not have perfect information.* If households are not aware of the prices and characteristics of the properties available in the market, then the implicit prices of property characteristics may vary from sale to sale and the HPF may be ill-defined.[19]

   Whilst it seems almost impossible for any one household to obtain perfect information it seems unlikely that the property market would support large-scale disparities in the prices of near-identical properties. For one thing, property markets in the UK tend to be well documented. Households can accumulate information on the characteristics and offer prices for properties in the market with relative ease and at practically no cost to themselves. Furthermore, the property market is highly mediated through the actions of estate agents and letting agents. These agents act on behalf of households and landlords providing as part of their service a (hopefully) in-depth knowledge of the pricing regime in the market.

2. *Transaction costs.* Transaction costs are the expenses, on top of the price of the property that the household incurs when moving house. In this context, for example, households' transaction costs would include the time spent searching for properties, expenses on lawyers and surveyors, taxes and the costs of moving possessions from one property to another.

   Transaction costs may prevent the market from reaching universal equilibrium. For certain households, the benefits of moving from their current property to a utility maximising residential location may be outweighed by the transaction costs. Since such households will not be at optimal residential locations, it can be argued that the market is in a state of disequilibrium.

   Again, such disequilibrium may not concern us a great deal. In particular, the applications presented in this thesis estimate HPFs using very large data sets

[19] Of course imperfect information characterises most market. Who hasn't walked into the first shop on the high street and bought the product in their range that best suits their purposes, then walked five minutes up the road to find a shop selling the same product at a cheaper price.

describing purchases in the property sales market. It seems reasonable to expect that at the moment of purchase households are trying to make utility maximising choices; even more so given the considerable transactions costs of moving. Likewise, we might expect that landlords are providing their properties to the market with the set of attributes that maximises their profit. Assuming that the number of households and landlords participating in the market is sufficient, one might assume that the reigning prices represent a market equilibrium for these particular economic agents.

3. *Slow adjustment of the hedonic price schedule.* For the market to reach equilibrium, changes in the conditions of supply or demand will be reflected in changes in the hedonic price schedule as households seek out their utility maximising residential location under the new conditions. In the real world, many issues including imperfect information and transaction costs will result in this process of adjustment taking some time.

   The key issue for empirical analysis is to collect data from periods of stasis in the market. In other words, satisfactory identification of the HPF is less likely when data are collected around the time of large-scale changes in property markets that may induce adjustments in the equilibrium price function. Furthermore, ensuring data is collected from a relatively narrow time frame should help avoid the possible impacts of more gradual changes in the price function.

At any one point in time, therefore, it seems unlikely that a property market will be in a state of universal equilibrium. However, provided certain precautions are taken in the collection of property market data, this does not present a major shortcoming of the hedonic price method.

A further concern discussed by Freeman (1993) is that the market may not contain sufficient variability in the combinations of characteristics available in properties supplied to the market. It is argued that for the market to attain equilibrium, households must be allowed to choose from all possible combinations of housing characteristics, since only then are they able to locate at a position of simultaneous equilibrium with respect to all characteristics.

An example of this problem was provided by Harrison and Rubinfeld (1978). Having estimated the HPF for housing in Boston, they were interested in calculating the marginal implicit price of air pollution. Contrary to expectation they discovered that high-income households were locating in regions of high air pollution. They suggest that an explanation for this phenomenon could be that some high-income households wished to locate in properties that provided both low levels of pollution and high levels of another attribute (e.g. ease of access to the cultural amenities found near the city centre). However, no properties satisfied both of these requirements. As such it is argued that certain households had to content themselves with a corner solution with respect to the level of air pollution. That is, the maximum utility they could attain from the properties available in the market was one in which their marginal WTP for improved air quality exceeded the implicit price of air quality.

That "gaps" exist in the range of products provided to the property market is undeniable. Freeman (1993, pp385-86) concludes that "the problem is almost certain to exist for some subgroups in some urban areas" and that "this is a problem to which empirical researchers must be sensitive".

However, whether the existence of such gaps is a serious concern for researchers is debatable. It seems clear to the author that the existence of gaps does not undermine the theoretical models described in Chapter 1. Rather, given the constraints imposed by the available housing stock, the market will settle on an equilibrium price function characterised by households choosing utility-maximising properties. That some of these households are at corner solutions is irrelevant to the establishment of a market-clearing price function.

Of course, the existence of gaps may cause practical problems in the estimation of the HPF. As is witnessed by the Harrison and Rubinfeld example described above, that problem is the familiar one of collinearity. If the set of properties available in the market do not provide sufficiently contrasting property attributes, it may be difficult to distinguish empirically the impacts of the different property characteristics on selling prices.

In contrast, the fact that households are at corner solutions may have serious impacts on attempts to econometrically identify the underlying structural equations (utility and cost functions) of the hedonic pricing model. We reserve discussion of such issues for Part 3 of this thesis.

3. Estimation of the Hedonic Price Function

3.i. Market Definition

In Chapter 1 of this thesis, we established that the hedonic price schedule for a property market is determined by the particular characteristics of the households and housing stock that make up that market. Since these factors will likely differ across property markets, the HPFs for different markets will themselves differ. A primary concern of researchers, therefore, will be to ensure that their data is drawn from a single market.

Hedonic studies have varied considerably in the geographical area that they have considered as one market. Some researchers have deemed it suitable to use data from house prices for an entire nation whilst, at the other extreme, some have focused on areas no bigger than a single census tract. Obviously it is possible to make one of two errors in collecting data for a hedonic study;

- Data is collected from only a small portion of a single property market
- Data is collected on house prices that come from more than one different property market

In the first case, there is a good chance that our estimates of the HPF will be imprecise. That is, data on only part of the market is unlikely to provide information on all the possible combinations and extremes of housing characteristics. Under such conditions, it will be difficult to define the true path of the hedonic price schedule with accuracy.

In the second case, we risk seriously biasing our estimates of the HPF. For example, imagine we collect property market data from two different markets that follow very different price schedules. If we were to estimate a single HPF based on the combined data from the two markets, the result will be a HPF that has no real world counterpart and is a poor reflection of both of the two underlying price schedules.

The key question facing researchers, therefore, is how to identify independent property markets. Theoretically two property markets will be independent if there exists one or more barriers preventing households and landlords in one market participating in another. Accordingly one might reasonably assume that two geographically separated cities are characterised by independent property markets. In this case, the costs of gathering information and physically moving between urban areas raise a geographical barrier that ensures the two property markets are independent.[20]

To test such an assumption, researchers have estimated a separate function for each suspected property market and tested to see whether the HPFs for the separate markets differ in statistically significant ways. Evidence from the literature using this sort of test reveals somewhat ambivalent results. For example, Butler (1980) tested to see whether a national housing market existed by comparing data from 36 cities in the USA. Though he concluded that the market in the sale and purchase of houses could not be considered a single market, he found it impossible to reject the possibility that the house rental market was a single national market. Smith and Huang (1995) surveyed hedonic pricing studies carried out between 1967 and 1988 and concluded that the estimated HPFs varied across cities due to differences in local conditions.

The key question for this thesis is whether data collected from a single urban area can reasonably be considered as emanating from a single market. To date, the hedonics literature seems somewhat confused on this issue.

Numerous authors have sort to partition data from a single property market along various dimensions under the assumption that the different partitions might represent different market segments (e.g. Straszheim, 1973, 1974; Ball and Kirwan, 1977; Schnare and Struyk, 1976; Sonstelie and Portney, 1980; Goodman, 1978; Michaels and Smith, 1990; Allen et al., 1995; Wolverton et al., 1999; Goodman and Thibodeau, 1998, 2003). For example, Basu and Thibodeau (1998) identify structure

[20] Of course, such a barrier is not insurmountable such that in the real world households do migrate between cities. Accordingly property markets of geographically separated urban areas may not be entirely independent, but for the purposes of this discussion this possibility is ignored.

type, structure characteristics and the characteristics of neighbourhoods as dimensions that might characterise such market segments.

In the main, these studies have compared the HPFs estimated from properties in the different partitions, concluding that market segmentation is a feature of an urban area if the HPFs for the separate partitions differ in statistically significant ways. Whilst a few isolated studies employing such tests have rejected the existence of market segmentation (e.g. Ball and Kirwan, 1977, found that HPFs estimated for different housing types in the Bristol area did not differ significantly from one another) the vast majority of studies have concluded that different market segments exist within single urban areas.

Besides this empirical evidence, a few authors have sought to provide a theoretical justification for partitioning data from one urban area. For example, Schnare and Struyk (1976) argue that segmentation will result whenever households' demand for a particular locational, structural or neighbourhood characteristic is highly inelastic and when this preference is shared by a relatively large number of other households. Under this line of reasoning one might justify partitioning property market data according to;

- *structure type*: households may wish to purchase a property of a certain type. For example, the market might segment between households looking to purchase houses with gardens and those looking to purchase flats or maisonettes.
- *structural characteristics*: households may have strong preferences for a particular property characteristic. For example, segmentation might result if certain households only consider buying period properties with "original features" whilst others only consider purchasing modern homes.
- *Neighourhood characteristics*: households may have strong preferences for localities providing certain amenities. For example, certain households may desire proximity to transport links or good quality schooling whilst others find no advantage in such proximity. Similarly, households may segment along income or racial lines particularly if households prefer to live in areas of relatively homogenous socio-economic characteristics.

In Chapter 5 of this thesis the Birmingham data set is partitioned according to a combination of property and neighbourhood characteristics. This partitioning defines easily interpretable market segments that are characterised by statistically different HPFs. Following Schnare and Struyk's (1976) justification these various partitions are interpreted as independent property market segments.

In Chapter 6, however, this justification is called in to question. In particular, it is argued that the technical definition of market segments requires that one or more *barriers* exist so as to prevent households and landlords in one market segment participating in another. As Palmquist (1991) points out, such market barriers may be geographical in nature, or may take the form of discrimination, or be precipitated by a lack of information. Clearly, heterogeneity in household preferences does not constitute such a market barrier.

Indeed, Chapter 6 puts forward an alternative explanation for the empirical observation that the HPF differs across partitions of the data. This explanation is based on the theoretical models of property market equilibrium described in Chapter 1 and Appendix A. In these models all households and landlords participate in the same market characterised by a single market clearing HPF. However, the model allows for the fact that heterogeneity in household preferences for property attributes may result in a highly nonlinear market clearing HPF. In this case households with different preferences may choose to locate at properties with very different *marginal* (implicit) prices for property attributes. Partitioning the data acts so as to isolate different regions of the hedonic price surface. Econometric estimation of the relationship between property prices and property attributes for each partition of the data simply provides a local approximation to the nonlinear HPF over that particular region of the hedonic price surface.

Discussion of these two distinct justifications for partitioning property market data is taken up in detail in Chapter 6.

3.ii. Changes in the HPF over time

As well as differing over space, one might expect that the HPF to differ over time reflecting changes in the conditions of supply and demand in the property market. Again, it is possible to test for temporal separation of markets using statistical techniques by comparing HPFs estimated from data for different periods.

Evidence from the hedonic literature suggests that temporal separation of markets may be a problem. Edmonds (1985) found that the HPFs estimated from two separate Japanese datasets from 1970 and 1975 were distinctly different. Palmquist (1980) found that the proposition that a dataset covering 13 years of house prices in Washington in the USA, represented information on one equilibrium hedonic price schedule was unacceptable. However, when adjacent pairs of years were used, the hedonic price schedule appeared to be reasonably stable.

Overall, it would seem wise to regard the aggregation of data over time with some caution. If the market has not been subject to any significant shocks during the period, aggregation may be defensible and statistical techniques can be used to test this hypothesis.

3.iii. The Dependent Variable

A primary concern in estimating a HPF is the measurement of property price. Early hedonic pricing studies in the United States used census data, which is relatively easy to obtain. As Freeman (1993) points out, however, there were some problems in that the data are presented in aggregate format that reduces accuracy and curtails the ability of the researcher to control for relevant housing and location characteristics.

A further problem with US census data is that it is on house prices based on home-owners' personal estimates. How closely these estimates reflect the true price that a property would command in the housing market is debatable. For example, Nelson (1978) showed that personal estimates of property values from US census data were 3% to 6% higher than those given by professional valuers.

This suggests a second possible source of data on property prices; professional valuations. It is not uncommon for large datasets to be compiled on the values of properties for the purposes of taxation. Again, data from these sources are not entirely reliable as they are, after all, only best guesses at the actual selling prices of properties (DoE, 1972).

Most commentators would agree, therefore, that by far the most preferred source of data for hedonic property market studies are records of actual sales prices on individual properties. Fortunately, such data is available for the studies discussed in the ensuing Chapters.

3.iv. The Explanatory Variables

Hedonic pricing is a data intensive technique. In the estimation of a HPF researchers must account for a very large number of explanatory variables. We would expect that the structural attributes of the accommodation itself, indicators of its accessibility, variables describing the characteristics of the neighbourhood and measures of environmental quality will all be important determinants of house price. These different categories of variables are summarised in Table 1.

Including accurate measurements of all the relevant explanatory variables in the specification of the HPF is extremely important. It is a fact of regression analysis that omitting or mismeasuring explanatory variables can lead to bias in the estimation of the regression parameters.[21]

Table 1: Categories and examples of variables in the hedonic price function

Variable Category	Examples of Variables in this Category
Structural	Number of rooms; presence of garage; size of garden; presence of central heating; etc.
Accessibility	Distance to: bus stop; town centre; school; shopping centre; etc.
Neighbourhood	Average age; race distribution; crime rate; quality of surrounding schools; etc
Environmental	Noise levels; air pollution levels; quality of views from the property; etc.

Source: Based on Tinch (1995)

Though it is usually relatively easy to define the relevant characteristics of the property itself (i.e. the *Structural* variables in Table 1), the definition and measurement of the other variables may be somewhat more complex.

Accessibility, for a start, is a rather vague concept. In hedonic pricing studies, a variety of accessibility measures are frequently included, for example distance to the CBD, access to main roads, distance to schools and distance to environmental facilities. Since accessibility variables are inherently spatial, recent developments in

---

[21] Chapter 7 concerns itself with such biases induced by the omission of variables relating to the geographical location of properties.

geographical information systems (GIS) have introduced much greater flexibility and precision into the estimation of accessibility variables. For example, it is possible to use GIS to calculate car travel times to important amenities that reflect the actual distance travelled on the road network taking account of road speeds along various road types. In the same way walking distances from a property to local amenities can be calculated precisely using the network of pedestrian routes (Lake et al., 1998).

Neighbourhood variables describe the characteristics of the local area in which the property is located. In general, census data are a good indicator of these attributes. Again GIS are a fast and efficient means of matching properties to census data at different spatial scales. For example, neighbourhood variables can be constructed from the smallest unit of the census, reflecting the direct neighbourhood of the property. Alternatively, user defined neighbourhood areas can be defined such as the area within 5 minutes walking distance of the property. GIS, therefore, allows researchers to efficiently and accurately consider the impacts of local as well as wider neighbourhood effects on the value of properties.

A final major concern is the definition and measurement of *environmental* variables. As several researchers have pointed out, the correct measure of an environmental variable will be the one that reflects households' perception of the environmental (dis)amenity. This is not as easy to obtain as it might sound. For example, a measurement of noise pollution should reflect many facets of this disamenity, including its intensity, frequency, duration, variability, time of occurrence during and so forth. Though we shall look at the issue in more detail in the next Section, there is clearly no single, obvious measure capable of reflecting all these aspects of noise pollution. It is important, therefore, that we consider carefully our measures of environmental variables.

To conclude, hedonic price studies must account for as many important explanatory variables as is possible. Failure to do so may lead to serious bias in the estimation of the parameters for the variables that are included. Fortunately, the ability of GIS to calculate large quantities of spatial data rapidly and accurately is a considerable technical advance in the compilation of explanatory variables. In Chapter 4 we describe the use of GIS to compile an extremely detailed property market dataset for the City of Birmingham in the UK.

3.v. Functional Form

A further important task facing researchers is to establish the exact nature of the relationship between the dependent variable and the explanatory variables; imposing an incorrect functional form on the regression equation will lead to *mis-specification bias*.

Unfortunately, economic theory gives no clear guidelines on how to select functional form. Typically, early example of hedonic price regression adopted one of four simple parametric functional forms;

- *Linear specification*; both the dependent and explanatory variables enter the regression in their linear form.
- *Semi-log specification*; the log of the dependent variable is regressed against linear explanatory variables
- *Log-linear specification*; a linear dependent variable is regressed against the log of the explanatory variables, and
- *Log-log specification*; both the dependent and explanatory variables enter the regression in their log form.

Whilst these simple parametric functional forms provide simplicity of estimation and interpretation they are highly restrictive. Since little in economic theory would support such strong assumptions, considerable attention has been focused on the use of more flexible specifications. In particular, a number of researchers have investigated the use of the Box-Cox flexible functional form (e.g. Cropper, Deck and McConnell, 1988; Cheshire and Sheppard, 1998). Whilst, this approach allows the regression model to more accurately reflect the patterns of association inherent in the data, it also has a number of drawbacks (as discussed by Ramussen and Zuehlke; 1990).

An alternative to increasing the degree of parameterisation of the regression model is to adopt a nonparametric regression approach. As is well known, however, nonparametric estimation is only realistic when there are only a small number of regressors. When there are many regressors, as tends to be typical of hedonic price studies, nonparametric response coefficients may be very imprecise.

An intermediate strategy is to employ a semiparametric form in which part of the model is specified parametrically whilst the rest is estimated using non-parametric techniques. Chapters 5 and 7 describe the application of semiparametric techniques to hedonic price estimation.

An alternative approach is developed in Chapter 6. Since the theoretical literature predicts that the equilibrium hedonic price surface will be highly nonlinear, it is argued that attempting to estimate a universally applicable function is impractical. Rather, a more realistic estimation strategy is to locally approximate the hedonic price surface over particular areas of property attribute space.

3.vi. Multicollinearity

A further problem that researchers often encounter in estimating the HPF is that of *multicollinearity*. Multicollinearity occurs frequently with environmental variables. For example, we would envisage that both noise pollution from traffic and air pollution from traffic fumes will have a negative impact on the price of a property. Unfortunately, the two are highly correlated. Higher levels of traffic result in greater noise pollution *and* higher concentrations of exhaust fumes. Regression analysis finds it extremely difficult to tease apart the separate influences on property prices of these two distinct but closely related disamenities of living in close proximity to a road.

Unfortunately, there is no easy solution to the problem of multicollinearity and in its presence estimated regression parameters may be implausibly large or in the worst case, have the wrong sign.

It is sometimes possible that problems of multicollinearity can be overcome through more accurate measurement of variables. For example, if data on noise pollution from roads takes account of local features such as trees and banks that act to dissipate traffic noise, then the correlation of noise pollution with air pollution may be less distinct. Once again, the power of GIS may be invaluable in this respect through improving the accuracy of variable measurement.

Another possible approach is simply to circumvent the problem by combining highly correlated variables into one index. This is the basis of a procedure known as *factor analysis* that is described in more detail in Chapter 5. In both that Chapter and

Chapter 6, factor analysis is used to combine variables describing the various socioeconomic characteristics of a property's neighbourhood into a more manageable set of readily interpretable indices.

3.vii. Spatial Autocorrelation

One final issue in the estimation of HPFs is that of *spatial autocorrelation* in regression residuals. By virtue of the fact that properties in close proximity share very similar environmental, accessibility and neighbourhood characteristics, property prices tend to be correlated over space. If it were possible to measure all these characteristics then spatial correlation in property prices could be accounted for in the explained part of the econometric model. Unfortunately, this is infrequently the case. Indeed, any empirical specification of the regressors in a HPF is unlikely to be sufficiently comprehensive to remove all spatial effects from the data. Of course, one can test this hypothesis by examining regression residuals for spatial autocorrelation. Evidence of positive spatial autocorrelation in regression residuals is an indication of spatial processes that are not captured by the specification of the HPF.

Unfortunately, traditional regression techniques such as Ordinary Least Squares (OLS) do not account for this correlation and researchers have had to adopt alternative estimation techniques. One such approach is to assume that spatial autocorrelation of the error terms results from the myriad subtle nuances of location that influence property prices. Under such an assumption one might adequately account for spatial autocorrelation by modelling the process generating. That is, to allow the error of each property to be functionally dependent on the errors associated with properties in its immediate neighbourhood. Such models have been widely applied in hedonic analysis (e.g. Dubin, 1988, 1992, 1998; Pace and Gilley, 1997; Basu and Thibodeau, 1998, 1999; Bell and Bockstael, 2000; Leggett and Bockstael, 2000). Indeed, in Chapter 5 just such a procedure is adopted.

Alternatively, spatial autocorrelation in regression residuals may reflect substantive spatial features whose absence from the model is likely to induce missing variable bias in the parameter estimates. For example, properties located close to an abattoir are likely to exhibit considerably deflated market prices. If proximity to abattoirs is not included as a regressor in the estimated HPF then one might conclude that the model is misspecified and that the parameter estimates are unreliable. In Chapter 7 we discuss in some detail a semiparametric estimation approach designed to account for such missing spatial variates.

4. The Valuation of Noise Pollution

4.i. Measuring Noise Pollution

In the last section we hinted at the difficulties of establishing a measure of noise pollution that truly reflected the impact noise has on households' lives. Broadly defined, noise can be described as unwanted sound and vibration (Litman, 1995). The qualifier "unwanted" covers a broad gamut of impacts that range from physiological (e.g. sleep disturbance), to pathological (i.e. auditory impairment) to psychological (Gent and Rietveld, 1993). The extent of these impacts will depend not only on the magnitude of the noise pollution but on its intensity, frequency, duration, variability and time of occurrence.

On top of this, the nature of noise pollution is truly multi-faceted. For a start vehicles produce noise from a whole variety of sources (e.g. mechanical movement, exhaust emission, tyre-road contact, aero-dynamic disturbances, bodywork vibration, brake friction, theft alarms, warning horns etc.). Furthermore, the level of noise pollution will depend on a number of parameters (Department of Transport, 1988) namely:

- Characteristics of the traffic itself; types of vehicles (e.g. share of motorbikes and HGVs), fluidity of traffic (closely related to the number of obstacles on the road such as traffic lights), traffic speed and drivers' behaviour;
- Characteristics of the road; type, width, state and quality;
- Characteristics of the environment: road side obstacles such as trees and fences, distance and height of receptor

A by-product of this complexity is that there is no 'correct' way to measure noise. Different noise exposure indices all try to convert the measured level of noise into a figure that reflects the perceived annoyance and hence take into account intensity, frequency, duration, variability, time of day and so forth (Nelson, 1978).

As a starting point the level of noise at any one point in time can be measured on a logarithmic scale using so-called 'A-weighted' decibels, dB(A). The dB(A) scale approximates the sensitivity of the human ear by weighting more heavily medium and high frequencies. Table 2 provides some examples and broad ratings of noise pollution using the dB(A) scale.

Table 2: The dB(A) scale with some examples

dB(A)	Rating	Sources	Effects
140		Gunshots, Explosions	Instant Auditive Trauma
130		Jet Aircraft Taking Off	
120			Pain above this Level
110		Pneumatic Drill	
100		Discotheque	
90		Close by a Lorry	
80		Busy Cross-Roads	Interference with Work
75	Very Bad		
70	Bad	Interior of Car	Interference with Speech
65	Quite Bad		
60	Moderate	Window on Busy Road	
55	Tolerable		
50	Quite Good	Quiet Street	Normal
45	Good		
40	Excellent	Calm Office	Interference with Sleep
30		Library	
20		Studio, Whispering	Sensation of Calm
10		Desert	
0			Limit of Audibility

Source: Adapted from Gent and Rietveld (1993), Soguel (1991) and Tinch (1995)

Clearly, a measure that only reflects the level of noise (usually represented by the letter L) at any one point in time is not an adequate reflection of the true diversity of noise pollution resulting from road traffic. An improved measure would be one that reflected the distribution of noise over the day. One possibility is to use the L_{10}, L_{50}

or L_{90} measures. These gauge the noise level exceeded 10%, 50% and 90% of the time and are known respectively as the peak, mean and ambient noise levels.

Alternatively, the equivalent continuous sound level or L_{eq} measure provides a single figure that reflects the distribution of sound throughout the day. Calculated in the same way as L_{eq}, the day/night equivalent sound level (L_{dn}) attempts to reflect the added annoyance of noise at night by weighting noise pollution at this time more heavily. A further measure is that of the noise pollution level (NPL). The NPL attempts to account for the added irritation of variability in the noise pollution experienced at any one location over the day. This is achieved by adding to the L_{eq} measure of noise pollution at a site a term that reflects the variability of noise levels.

4.ii. Studies of Noise Pollution

A large number of hedonic price studies have investigated the impact of noise pollution on property prices. In order to facilitate comparison of the results of hedonic price studies researchers often quote a Noise Sensitivity Depreciation Index (NSDI). Originally introduced by Walters (1975), the NSDI was adopted for comparative purposes by Nelson (1980, 1982) in his major reviews of hedonic price studies. For two residential properties that differ only in their level of noise exposure, the house price depreciation per dB can be defined as;

$$D = \frac{reduction\ in\ property\ value\ from\ noise\ exposure}{difference\ in\ noise\ exposure}$$

NSDI is then defined as;

$$NSDI = \frac{D}{property\ value} \times 100 = \frac{\%\ depreciation\ in\ property\ price}{difference\ in\ noise\ exposure}$$

NSDI can be regarded as a percentage change in price arising from a unit increase in noise. NSDI for the various studies is shown in the fifth column of Table 3. Values in the studies listed range from 0.08 (Palmquist, 1980 and 1981) to 2.22 (Gamble et al., 1974). The simple mean for these studies is an NSDI of around 0.55. In other words, these studies suggest that an increase in noise pollution of 1 dB(A) will reduce the value of a property by just over ½ of one percent. A list of hedonic pricing studies that have identified the NSDI for road traffic noise is provided in Table 3.

Table 3: Hedonic pricing studies of loss in property value from *Road Traffic* noise (% depreciation in house prices per 1 dB(A) increase in noise level)

Source Study	Study Year	Study Area	Noise Measure	NSDI
Allen, 1980†	1977-79	North Virginia, Va., USA	L_{10}	0.15
	1977-79	Tidewater, Va., USA	L_{10}	0.14
Anderson and Wise, 1977†	1969-71	Towson, Md., USA.	NPL	0.43
	1969-71	North Springfield, Va., USA	NPL	0.14
Bailey, 1977†	1968-76	North Springfield, Va., USA	Log of Distance	0.3
Gamble et al., 1974†	1969-71	Bogotoa, N.J., USA	NPL	2.22
	1969-71	Rosendale, Md., USA	NPL	0.24
	1969-71	North Springfield, Va., USA	NPL	0.21
	1969-71	All three areas	NPL	0.26
Grue et al.,1997		Oslo, Norway – *Obos*	L_{eq}	0.24
		Oslo, Norway – *Flats*	L_{eq}	0.21
		Oslo, Norway – *Houses*	L_{eq}	0.54
Hidano et al., 1992*		Tokyo, Japan	L_{eq}	0.7
Hall et al., 1978†	1975-77	Toronto, Canada	L_{eq}	1.05
Hall et al., 1982		Toronto, Canada – *Arterial*	L_{eq}	0.42
		Toronto, Canada – *Expressway*	L_{eq}	0.52
Hammar, 1974		Stockholm, Sweden	L_{eq}	0.8 –1.7
Iten and Maggi, 1990		Zurich, Switzerland	-	0.9
Langley, 1976†	1962-72	North Springfield, Va., USA	NPL	0.22
Nelson, 1978†	1970	Washington, D.C., USA	L_{dn}	0.87
Palmquist, 1980, 1981†	1962-76	Kingsgate, Wa., USA	L_{10}	0.48
	1958-76	North King County, Wa., USA	L_{10}	0.3
	1950-78	Spokane, Wa., USA	L_{10}	0.08
Pommerherne, 1988	1986	Basel, Switzerland	L_{eq}	1.26
Renew, 1996a		Brisbane, Australia	L_{eq}	1.0
Soguel, 1991	1990	Neuchatel, Switzerland	L_{eq}	0.91
Vainio, 1995		Helsinki, Finland	L_{eq}	0.36
Vaughan & Huckins, 1975†	1971-72	Chicago, USA	L_{eq}	0.65

† Reviewed in Nelson (1982)

* From Bertrand (1997) who notes that figure is presented with caution

The variety of NSDI values should not come as a surprise. Theoretically, we would not expect different housing markets to have the same HPF and, therefore, would not expect applications of the hedonic pricing technique in different cities in different years to return identical results.

Bertrand (1997) presents a meta-analysis drawing on 16 estimates from nine different hedonic pricing studies of noise pollution carried out in the USA, Canada, Switzerland and Finland. Extrapolating his results, the average NSDI in this selection of studies is found to be 0.64%. Bertrand also finds that the greater the average level of noise in a market and the greater the income of the market's households, then the higher the implicit price that is paid for noise pollution reductions. Whilst this might seem intuitively plausible, there is little theoretical support for the existence of such relationships (see Section 5 of Chapter 8).

The use of a single statistic to compare studies conceals considerable heterogeneity in the exact method of their application. For example, the studies vary with regards to the measure of noise used in the analysis (see Column 4 of Table 3). Likewise, the method by which the noise pollution impacting on a particular house is assessed differs from study to study. Some studies construct noise contours across the urban area by extrapolating from various monitoring points. The noise pollution experienced by any particular property will depend on the band in which it falls (e.g. Gamble et al., 1974; Palmquist, 1992). More advanced measures of noise pollution can be achieved by using models that take account of the exact characteristics of a particular dwelling (e.g. Pommerehne, 1987; Soguel, 1991; Vainio, 1995).

Studies, also vary considerably in the choice and accuracy of the *explanatory variables* used in the regression analysis and in the choice of *functional form*. For example, Vaughan and Huckins (1975) in an hedonic price study using individual housing sales in the Chicago urban and suburban areas in 1971-72, included variables reflecting structural characteristics (e.g. sq. feet of living space, number of garages, lot width and age of dwelling), neighbourhood characteristics (e.g. total number of lots on the block – a crowding measure, number of visible broken windows – a blight measure and available recreation land), accessibility characteristics (e.g. distance to the central business district and distance to Lake Michigan) and environmental characteristics (i.e. noise pollution and air pollution – a composite measure of sulphates and particulates). In contrast, Allen (1980) again using individual housing sales but this time for two towns in Virginia only included structural variables (e.g. sq. feet of floor space, sq. feet of lot, no. of bathrooms, no. of fireplaces, age of property etc.) as well as a measure of noise pollution.

A number of hedonic studies have focused on the impacts of aircraft noise on property prices. A list of such studies is provided in Table 4. These display a variety of NSDI scores ranging from 0.29 up to 2.3. The mean NSDI score for these studies is 0.87 though this falls to 0.64 when the relatively high figures reported by the Paik (1972) and Yamaguchi (1996) papers are removed. This is roughly in line with Nelson's (1980) conclusion that "... the noise discount is commonly 0.5-0.6% ...".

Schipper (1996) has carried out a more formal statistical test of these results using meta-analysis. He finds that the implicit price of aircraft noise pollution is influenced by a number of factors including the timing, country and specification of the original noise studies. His findings suggest that as a baseline the NSDI is around 0.33% whilst for studies in the United States this rises to 0.65%. Again, there is little theoretical support for studies of this kind.

Only Pennington et al. (1990) failed to return a significant and negative coefficient on the noise variable. Pennington et al. (1990) undertook a hedonic price study using property data of actual market transactions covering the period April 1985 to March 1986. They found that, following the inclusion of neighbourhood variables, the noise variable became statistically insignificant. As such they concluded that differences in property values could be attributed solely to neighbourhood and other characteristics.

The Pennington et al. (1990) result was challenged by Collins and Evans (1994). They studied the noise effect of Manchester Airport with the same data set but using a non-regression analytical technique using artificial neural networks (ANN). Though it is impossible to judge the significance of coefficients using the ANN approach, Collins and Evans distinguished a sizeable noise effect despite the fact that this was dwarfed in importance by the impact of neighbourhood characteristics.

Concerning *functional form*, the majority of researchers have opted for the 'traditional' semi-log form. However, there is no theoretical reason to believe that this is the optimal specification, indeed, Levesque (1994) employed the Box-Cox transformation and showed that the functional form implied by his data was significantly different from the semi-log.

Table 4: Hedonic pricing studies of loss in property value from *Aircraft Noise* (% depreciation in house prices per 1 dB(A) increase in noise level)

Source Study	Study Year	Study Area	NSDI
Abelson, 1979†	1972-73	Marrickville, Sydney, Australia	0.4
	1972-73	Rockdale, Sydney, Australia	0.5
Collins and Evans, 1994	1985	Manchester, UK	-
De Vany, 1976†	1970	Dallas, USA	0.8
Dygert, 1976†	1970	San Mateo, San Francisco, USA	0.5
	1970	Santa Clara, San Jose, USA	0.7
Emerson 1969, 1972†	1967	Minneapolis	0.58
Gautrin, 1975†	1968-69	London Heathrow, UK	0.62
Levesque, 1994		Winnipeg, USA	1.3
McMillan et al., 1980†	1975	Edmonton, Canada	0.51
Maser et al., 1977†	1971	Rochester, N.Y., USA – *City*	0.88
	1971	Rochester, N.Y., USA – *Suburban*	0.61
Mieskowski & Saper, 1978†	1969-73	Etobicoke, Toronto, Canada	0.52
Nelson, 1978†	1970	Washington, USA	1.06
Nelson, 1979	1970	San Francisco, USA	0.58
	1970	St. Louis, USA	0.51
	1970	Cleveland, USA	0.29
	1970	New Orleans, USA	0.4
	1970	San Diego, USA	0.75
	1970	Buffalo, USA	0.52
O'Byrne et al., 1985	1980	Atlanta, USA	0.69
	1970	Atlanta, USA	0.64
Paik, 1972†	1960	New York, USA	1.9
	1960	Los Angeles, USA	1.8
	1960	Dallas, USA	2.3
Pennington et al., 1990	1985	Manchester, UK	0.47
Price, 1974†	1960-70	Boston, USA	0.83
Uyeno et al., 1993	1987	Vancouver, Canada	0.65
Yamaguchi, 1996	1996	London Heathrow, UK	1.51
	1996	London Gatwick, UK	2.30

† Reviewed in Nelson (1980)

As for the use of *explanatory variables*, all the studies contain structural characteristics whilst treatments of neighbourhood and accessibility characteristics differ from one study to another. For example, the Levesque (1994) study did not include neighbourhood characteristics, whilst, as already mentioned, the Pennington et al. (1990) study claimed that including neighbourhood characteristics made the noise coefficient insignificant. On the other hand, O'Byrne et al. (1985), studying the impact of noise pollution from Atlanta International Airport, obtained more or less the same noise effect with and without neighbourhood characteristics included in the regression.

With regard to accessibility characteristics, three of the aircraft noise studies (Levesque, 1994; Nelson, 1980 and Uyeno et al., 1993) included these variables, each confirming the importance of including accessibility variables in hedonic pricing regressions. Both Nelson (1980) and Uyeno et al. (1993) attempted to account for the importance of accessibility to the airport as a focal point for employment, transportation and commercial services. However, the latter researchers dropped this variable from their final specification because they found it an insignificant factor in determining property prices. Conversely, Nelson found this accessibility variable to be significant and concluded that major bias could be introduced into hedonic pricing studies if it is ignored.

4.iii. Noise Pollution and other Valuation Approaches

Confirmation of the robustness of the hedonic pricing method can be sought through comparison with figures derived from other valuation methodologies. For example, JMP Consultants Ltd. (1996) carried out research for the Department of Transport valuing the nuisance from road traffic. One approach they adopted (and in their opinion "the most plausible and practical tool for the valuation of nuisance arising from changes across a broad spectrum" p 256.) is to ask the opinion of expert property valuers. Using a large sample they concluded that the best estimate of the NSDI was 0.29% per dB increase or decrease in noise pollution. This result falls in the range of values commonly reported from hedonic studies but is somewhat lower than the average of values reported in the hedonic literature.

A further point of comparison can be found in studies that have employed more than one methodology to investigate a single valuation problem. For example, Pommerehne (1987), Soguel (1991 and 1994) and Vainio (1995) have used the contingent valuation approach to produce results that they compare with those derived from their hedonic analyses.

The Pommerehne (1987) study in Basel, Switzerland produces remarkably similar results. Estimating households' WTP to reduce noise pollution by half, the hedonic price method returns a result of 79 Swiss Fr per month compared to a value of 75 Swiss Fr per month derived from the contingent valuation survey.

Similarly the Soguel study in Neuchatel, Switzerland produces highly similar results. Again valuing households' WTP to reduce noise pollution by half, the research estimates a value of 60 Swiss Fr per month from the hedonic pricing method (Soguel, 1991) and a value of between 56 and 67 Swiss Fr per month from the contingent valuation method (Soguel, 1996).

In contrast, the Vainio (1995) study in Helsinki, Finland concludes that a change in noise pollution levels from L_{eq} 65 to L_{eq} 55 would be valued at FIM 18,420 using the hedonic pricing method and at almost three times this amount (FIM 51,600) using the contingent valuation approach.

Of course, a great deal of caution should be taken in making such comparisons. Contingent valuation returns estimates of households' maximum WTP (WTP) for changes in noise exposure. Hedonic pricing determines the price that must be paid in the market for such changes. Clearly the two techniques are not measuring the same quantities. Indeed, Part 3 of this thesis is dedicated to determining how information from hedonic markets might be used to provide estimates of WTP.

4.iv. Summary

The evidence presented in this section provides considerable support for the contention that noise pollution is capitalised into property prices. The estimated impact of noise on house prices would appear to vary considerably from study to study. This is not surprising considering studies are taken from different markets and at different times and, therefore, will be estimating quite separate equilibrium HPFs.

The values quoted as NSDIs range from 0.08% to 2.22% for road traffic noise and from 0.29% to 2.3% for aircraft noise. Statistical analyses of these results suggest average values of 0.64% for road traffic pollution and 0.33% for air traffic pollution (or 0.65% in the United States).

5. Summary and Conclusions

This chapter has provided a basic overview of theoretical and empirical considerations associated with the statistical estimation of the HPF. Several conclusions can be drawn from this discussion.

First and foremost, it would appear that, provided we are careful in our selection of data, the assumptions of the hedonic price model are not so unrealistic as to make our estimations of the hedonic price schedule from real world data meaningless.

Second, as far as the actual estimation of the hedonic price schedule is concerned, a number of possible problems and pitfalls have been highlighted. However, given careful consideration and use of state of the art techniques, including;

- the use of GIS to calculate accurate explanatory variables
- the partitioning of the data into market segments
- the employment of flexible functional forms and
- allowing for spatial autocorrelation

it would appear that none of these problems are insurmountable.

As the review in Section 4 has shown, a large weight of evidence has now been amassed to support the contention that environmental quality, such as exposure to noise pollution, is capitalised in the price of property. In a very large number of studies, the hedonic pricing method has been successfully employed to identify the size of this impact.

CHAPTER 4: THE CITY OF BIRMINGHAM PROPERTY MARKET DATA SET

1. Introduction

The case study that forms the core of this thesis concerns the City of Birmingham property market. Whilst our eventual goal is the estimation of the HPF for the City of Birmingham, this Chapter concentrates on describing the collation and pre-processing of the data set.

As described in Sections 2 and 3 of this Chapter, the data for this study have been drawn from a wide variety of sources; some in the form of electronic data bases, others on sheets of paper tucked away in ranks of filing cabinets. The data have been constructed and collated using geographical information systems (GIS), a powerful tool designed to deal with the large quantities of spatially referenced data. The resulting data set is by far the richest of its kind in the UK and almost certainly the largest and most comprehensive hedonic data set yet compiled anywhere in the world.

Unfortunately, the quantity of information provided by the application of GIS complicates the estimation of the HPF. In particular, multicollinearity is rife in the covariate data, with many variables measuring slightly different dimensions of the same basic property characteristic. To overcome this difficulty, Section 4 describes the application of factor analysis to the covariate data. Factor analysis provides a way in which the multitude of variables available to the analyst can be concentrated into a smaller number of factors that identify the major dimensions of difference and similarity between properties. Summary statistics for the final data set used in the regression analysis reported in the next Chapter are provided in Section 5.

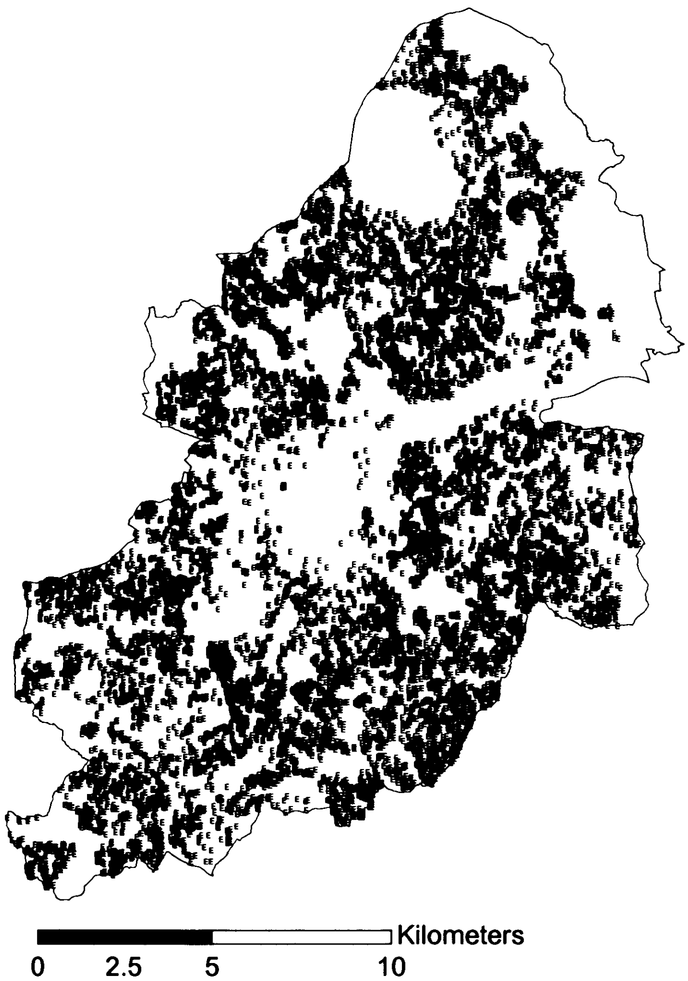
As indicated in Chapter 3, one of the primary innovations of the work presented in this thesis is in the use of model-based clustering techniques to partition the data into groups of properties displaying similar characteristics. In the application discussed in Section 6 of this Chapter, properties are grouped according to both their structural attributes and also according to the socioeconomic characteristics of their neighbourhoods. Model-based clustering represents the cutting edge in cluster analysis. As far as the author is aware these techniques have not previously been employed in hedonic analysis or in any other field of empirical economics. In this Chapter, in contrast to Chapter 6, groupings of similar properties identified by the cluster analysis are interpreted as potential property market segments. Section 6 provides provides a brief introduction to the techniques of model-based clustering, discusses the application of these techniques to the Birmingham data set and provides descriptions of the assumed market segments identified by the analysis.

2. The Dependent Variable

The preferred source of data for hedonic property market studies are records of actual sales prices on individual properties. Unfortunately, such data are not publicly accessible information in England and Wales. For this study the DETR (now DfT) arranged access to the databases of the UK Land Registry which maintains computerised records of all property transactions within the UK. Records of all property sales within the administrative boundaries of the City of Birmingham during 1997 were obtained indicating selling prices, dates of sales and full property address for each residential property transaction. All commercial properties and land-only transactions were excluded from the search. Further first-time right-to-buys were ignored since it was assumed that these would not reflect the full market price of the property being purchased. Finally, only houses were included in the final data set. Flats and maisonettes were omitted for two reasons; first, because they represent substantially different residential formats that might well be expected to trade in a separate submarket to houses. Second, because without knowledge of the vertical location of such properties it was almost impossible to determine their exposure to transport-related traffic noise.

Each property address was matched to an entry in the OS ADDRESS-POINT database, which provides a unique grid reference for each postal address in the UK (Ordnance Survey; 1996). Using these grid references the properties were located on a digital outline of Birmingham and allocated to a buildings outline on OS Land-Line.Plus (Ordnance Survey, 1996). The resulting data set contained 13,547 records of property transactions, the locations of which are illustrated in Figure 1.

Figure 1: Location of properties in the Birmingham City study sample



3. The Explanatory Variables

This section discusses the compilation of the data set and the selection of variables included in the specification of the HPF. To frame our discussion we follow the data classification provided in Table 1 of Chapter 3.

3.i. Structural Characteristics

The structural characteristics of properties in the dataset were gathered from two sources; from records kept by the Valuation Office Agency (VOA) and from the digital map data provided by Land-Line.Plus.

The Valuation Office Agency is an executive agency of the Inland Revenue, one of whose main functions is to value property for Council tax purposes. In order to perform this function, the VOA maintain a database describing the structural characteristics of every residential property in England. This is the first hedonic price study in the UK to make use of VOA data.

Unfortunately, the VOA data sources are currently held as paper records. As such the most up-to-date information for each property were obtained by manually searching through VOA files and recording details in a spreadsheet application.

As detailed in Table 1, the data collected from the VOA provided the basic structural characteristics of each property. Furthermore, the VOA classifies properties according to age and style of construction into one of around 30 property types called Beacon Groups. This information was also recorded as it provides a useful additional indication of property quality that cannot be determined from size and age alone.

Table 1: Structural variables obtained from the VOA

Variable	Description
Beacon group	33 nationally defined property groups defined by the VOA. These identify similar properties in terms of style and age
Floor area	Floor area in square metres.
Property type	20 property types (e.g.detached, semi-detached)
Property age	7 age bands with properties built after 1973 coded with the actual year of construction
Number of bedrooms	Number of bedrooms at each property
Number of WCs	Number of internal WC's at each property
Central heating	Central heating classification; full, partial or none
Central heating type	Central heating type; recorded as either radiators, warm air or night storage heaters
Double glazing	Double glazing classification; full, partial or absent
Garage	Garage classification; coded as either single, double, car port or none.
Parking	Other parking facilities recorded as car space, shared drive or rear entry
For further information see: Dwellinghouse Coding; an illustrated guide, The Valuation Office	

Using the Land-Line.Plus digital map automatic procedures were developed to define and extract the outline of each property's plot and buildings. Each property was then visually inspected using the GIS, in order to ensure that the building and plot area had been properly delineated. Subsequently, measures of ground floor area and garden were calculated for each property.

Descriptions of the *structural characteristics* of properties included in the specification of the HPF are to be found in Table 2. Also included in that table are the researcher's *a priori* expectations concerning each variables impact on property prices.

Table 2: Structural variables included in the Hedonic Price Models

Variable	Code	Description and *a priori* Expectations
Floor Area (m^2)	Area	Larger properties command higher prices
Garden Area (m^2)	Garden	Properties with larger gardens command higher prices
Number of Bedrooms (dummy vars.)	Bedroom 1 ⋮ Bedroom 12	Properties in the data set have between 1 and 12 bedrooms. We create 12 constants one for each number of bedrooms. 3 bedrooms is taken as the baseline and this constant is dropped from the regressions. Properties with more bedrooms command higher prices.
Number of WCs (dummy vars.)	WC 1 ⋮ WC 5	Properties in the data set have between 1 and 5 WCs. We create 5 constants, one specific to each number of WCs. One WC is taken as the baseline and this constant is dropped from the regressions. Properties with more WCs command higher prices.
Number of Storeys (dummy vars.)	Storey 2 ⋮ Storey 7	Properties in the data set have between 1 and 7 storeys. We create 6 constants specific to each number of stories between 2 and 7 inclusive. We include a separately labelled set of indicator variables for bungalows and set 2 storeys as the baseline category. Given that two properties have the same floor area it is expected that those with less storeys will be preferred to those with more storeys.
Garage (dummy var.)	Garage	Indicator variable identifying properties with a garage. Properties with a garage command higher prices.
Central Heating (dummy var.)	Central Heating	Indicator variable identifying properties with a central heating. Properties with central heating will command higher prices.
Age of property (decades)	Age	Age of the property in decades prior to 1997. The relationship between property age and property price is not entirely clear. Older properties may be desired for their "character" and "original features", more modern properties for their state of repair and more up-to-date facilities. What is clear, however, is that property age proxies for a number of property characteristics not least of which will be the architectural design of the house.
Property Type (dummy vars.)	Detached House Semi-Detached House	Semi-detached houses are taken as the baseline property type since all market segments contain properties of this

Variable	Code	Description and *a priori* Expectations
	End Terrace House Terrace House Detached Bungalow Semi-Detached Bungalow End Terrace Bungalow Terrace Bungalow	type. The coefficients estimated on the other property type dummy variables, reflect the relative difference in price between that property type and a semi-detached house with exactly the same characteristics. It is expected that houses will fetch more than bungalows. Moreover, properties will increase in value from terraces through end terraces and semi-detached properties through to detached properties.
Beacon Group (dummy vars.)	BG 1 (Unrenovated cottage pre 1919) BG 2 (Renovated cottage pre 1919) BG 3 (Small "industrial" pre 1919) BG 4 (Medium "industrial" pre 1919) BG 5 (Large terrace pre 1919) BG 8 (Small "villa" pre 1919) BG 9 (Large "villas" pre 1919) BG 10 (Large detached pre 1919) BG 19 (Houses 1908 to 1930) BG 20 (Subsidy houses 1920s & 30s) BG 21 (Standard houses 1919 to 1945) BG 24 (Large houses 1919 to 1945) BG 25 (Individual houses 1919 to 1945) BG 30 (Standard houses 1945 to 1953) BG 31 (Standard houses post 1953) BG 32 (Large houses post 1953) BG 35 (Individual houses post 1945) BG 36 ("Town Houses" post 1950)	In the models, BG 21 (standard houses built between the war) is taken as the baseline beacon group since all market segments contain properties of this type. The coefficients estimated on the other beacon group dummy variables, reflect the relative difference in price between that properties of that beacon group and a property in beacon group 21 with similar characteristics. The beacon group data collected from the VOA provides a detailed categorisation of properties according to their age, size, architectural type and quality. As such, we would expect these dummy variables to be important descriptors that add significantly to the explanatory power of the model.

To introduce the maximum flexibility into the specification of the hedonic function a series of dummy variables have been created to represent each different quantity of bedrooms, each different quantity of WCs and each different number of storeys. One important omission from the model is an indicator of whether a property had double-glazing. Unfortunately, this information was not recorded by the VOA for every property and was considered too unreliable to include in the analysis.

The variables listed in Table 2 provide a comprehensive description of each property's structural characteristics. Indeed, with reference to structural variables alone, the Birmingham dataset rivals all other published hedonic datasets.

3.ii. Neighbourhood Characteristics

Details of the characteristics of the inhabitants of neighbourhoods were extracted from the 1991 UK census. This data set is obtained by the Office for National Statistics (ONS) by surveying all households and publishing the results by grouping contiguous properties together to preserve confidentiality (Openshaw, 1995). The smallest scale at which census data is released is that of the enumeration district (ED). In Birmingham, on average, this consists of data for 191 households.

The census provides a myriad of information on the socioeconomic characteristics of the population living in each ED. Certain of these were selected for use in the analysis and these are described in Table 3. As described in detail in Section 4, we apply techniques of factor analysis to condense this myriad of variables into a more manageable set of indices that pick out the essential dimensions along which neighbourhoods differ in their socioeconomic.

Table 3: Enumerator District attributes collated from the 1991 census

Attribute	Attribute Description
No car	% households with no access to a car
Two cars	% two car households
Unemployment	% unemployment
Non-owners	% residents not owning their home

One-parent families	% lone parent households
Low Social Class	% residents in lower social classes
Families	% households with children
Age 0 to 10	% residents less than 10 years of age
Age 11 to 17	% residents aged 11 to 17
Age 18 to 24	% residents aged 18 to 24
Age 25 to 34	% residents aged 25 to 34
Age 35 to 49	% residents aged 35 to 49
Age 50 to 64	% residents aged 50 to 64
Age > 65	% residents over the age of 65
Over Crowding	% households with > 1 person per room
Non White	% ethnically non-white residents
Black	% ethnically black (African or Caribbean) residents
Asian	% ethnically Asian residents

3.iii. Accessibility Characteristics

Accessibility variables define the ease with which people can travel to local amenities (and disamenities) and can also provide information about the "quality" of those amenities. As described in Table 4 information on the location of a wide range of local (dis)amenities were collected from a number of different data sources.

The data on accessibility were constructed with the use of GIS. Since measurements of accessibility should relate to people's perception of distance, several measures were constructed each of which might relate to people's perception of distance. As illustrated in Figure 2 GIS were used to develop three different measures of accessibility; straight-line distance, car travel time and walking distance.

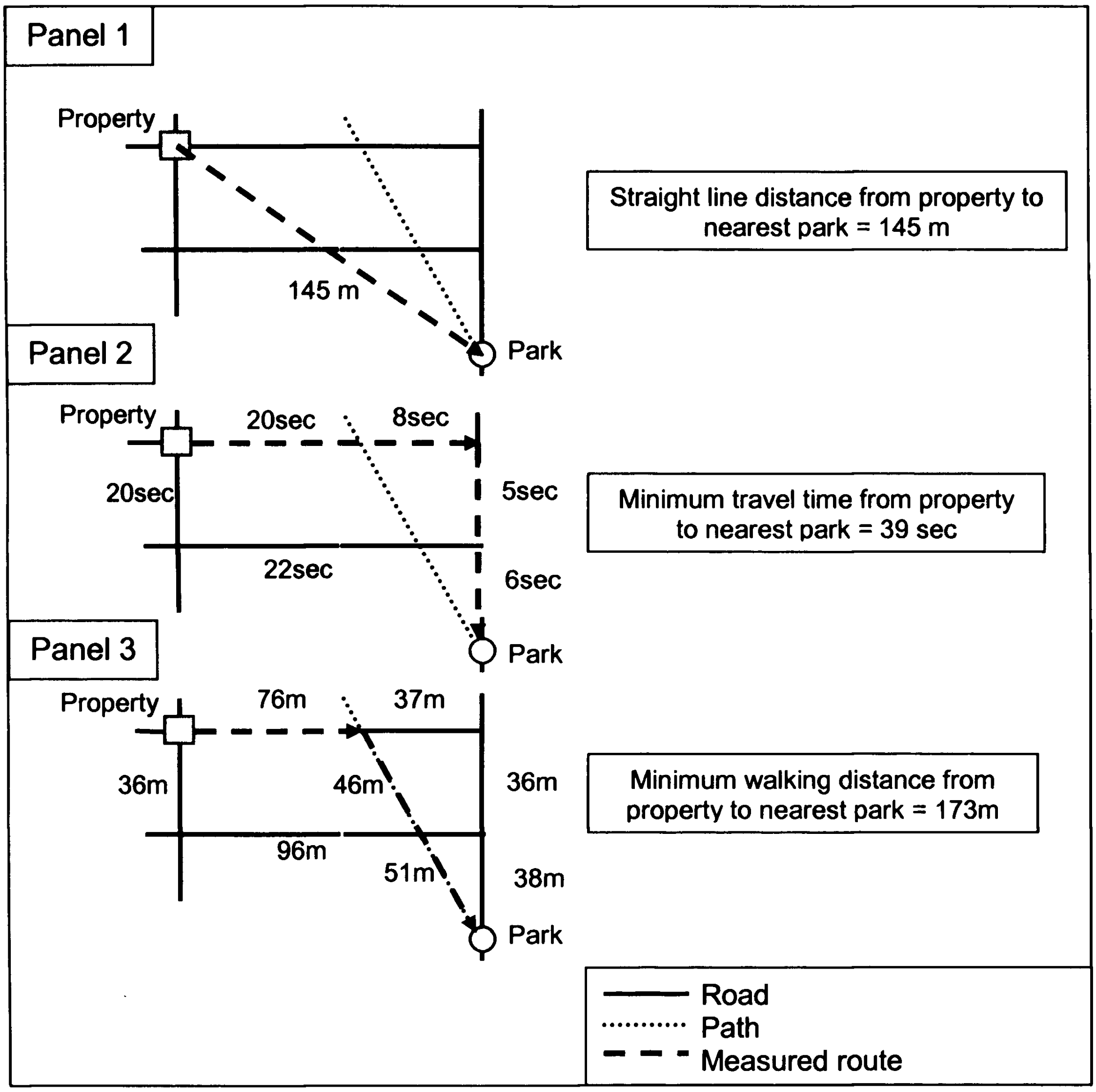
Table 4: Sources of amenity information

Amenity	Source of Information
Public parks	Any sites with the name "park", "recreation ground" or "common" were identified from the Birmingham A-Z and their outline extracted from Land-Line.Plus
Central Business District	Located as the Bull Ring in the centre of Birmingham
Birmingham Airport	Located as the B4438 at the Birmingham Airport Turn off
Railway stations	Locations obtained from Railtrack plc. and grid referenced using Land-Line.Plus
Primary Schools	Names and addresses determined from the OFFSTED web site and grid referenced using ADDRESS-POINT
Secondary Schools	As above.
Shops & groceries	Any businesses registered as "Delicatessens", "Grocers", "Newsagents" or "Supermarkets" were obtained from the Yellow Point database. Address locations were confirmed with ADDRESS-POINT to provide an accurate grid reference.
Industrial Sites	Locations of all Type A and B industrial process and landfill sites were obtained from Birmingham City Council.

Panel 1 of Figure 2 illustrates the basic measure of straight-line distance between each house and an amenity. Panel 2 illustrates accessibility measures based on car travel times. These are calculated from the road network provided by Land.Line.Plus. For each section of road, the time that a car would take to travel along its length is estimated based on the length of the road and the likely speed of the vehicle (road speeds being based upon DOT, 1993). The shortest travel time between each property and the nearest example of each amenity type is then calculated. Walking distance is calculated by using a road network, amended to exclude routes unsuitable for pedestrians (e.g. motorways) and supplemented with pedestrian only routes such as paths in parks and road cul-de-sacs through which pedestrian movement is possible. The distance along each section of this network was calculated and the shortest walking distance from each property to its nearest example of each amenity determined (Panel 3).

Using the GIS it would be possible to construct a myriad other variables including proximity to other types of retail outlet, proximity to major thoroughfares and proximity to bus routes. Whilst each of these locational characteristics might influence property prices, they are all essentially different, though similar indicators of the proximity of the property to local commercial centres.

Figure 2: Calculating accessibility variables

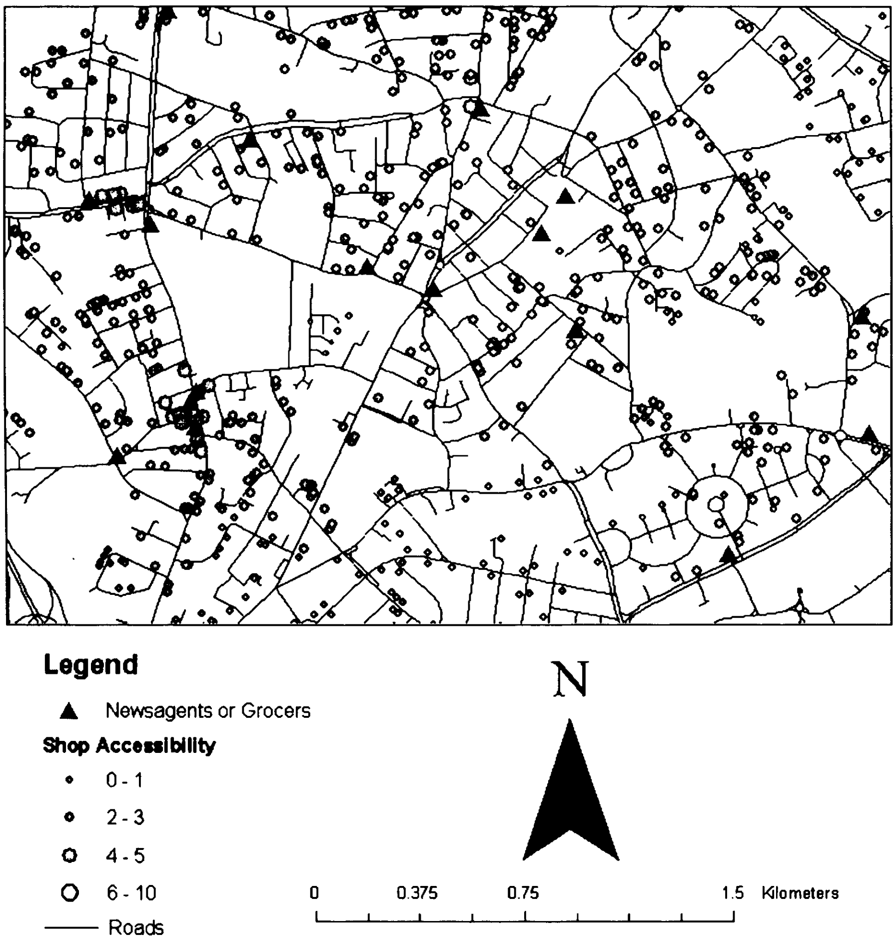


As such, a measure was required that captured both the proximity of a property to local centres and the relative size of those local centres. We use shops and grocers (see Table 4) as indicators of local commercial centres. Larger centres will have more such stores. Properties closer to local commercial centres will have better accessibility. Hence, our indicator of access to local commercial centres is constructed using a weighted sum of walking distances to all shops according to a common procedure in accessibility studies formalising to:

$$A_i = \sum_{j=1}^{J} \alpha_j e^{-\delta d_{ij}} \qquad (1)$$

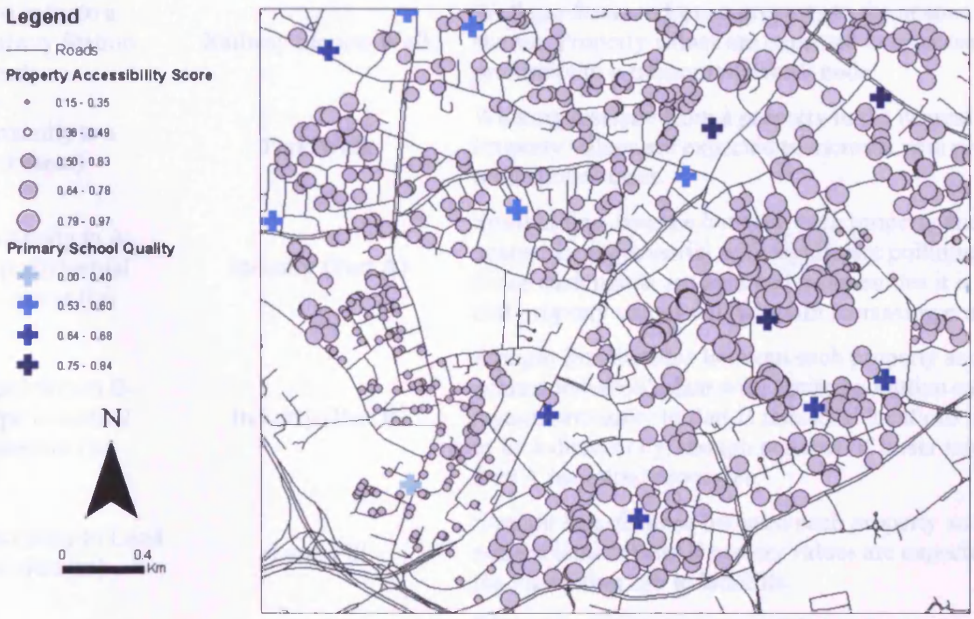
Where, A_i is accessibility at property i, α_j is the attractiveness of shop j, d_{ij} is the walking distance in kilometres between property i and shop j, δ exponent for distance decay and J is the number of shops in the region. Here we set $\delta = 2$ (such that a shop 100m from the property receives a weight over 6 times that of a shop at 1km distance and shops at over 2km distance receive almost no weight at all) and $\alpha_j = \alpha = 1$ (such that all shops are considered equally attractive). This shop accessibility variable is illustrated in Figure 3.

Figure 3: Shop accessibility scores for a selection of properties in the data set



A similar procedure was used when considering accessibility to primary schools. Recent research suggests that selection procedures for primary school intake that favour local residents can considerably inflate house prices around high performing schools (Gibbons and Machin, 2001).[22] For each primary school in the Birmingham area an estimate of school quality was calculated as the percentage of pupils achieving Level 4 or above in Science, Mathematics and English (the level expected of 11 year olds).[23] An accessibility index was constructed using (1) with the weight α_j set to this measure of school quality and $\delta = 1$. Figure 4 presents the primary school quality/accessibility variable depicted for a region of the study area.

Figure 4: Primary school accessibility scores for a selection of properties in the data set



Descriptions of the locational variables included in the specification of the HPF are provided in Table 5.

---

[22] The issue is thought less important for secondary schools that typically draw from much wider catchments (Gibbons and Machin, 2001). Also, high educational achievement at primary school level may be a pre-requisite for admission to selective secondary schools. For example, the five selective Grammar Schools of King Edward the Sixth in Birmingham make offers "… solely on the basis of performance in the entrance test. Special allowances are not made for brothers or sisters or distance from the school." (quote taken from the Grammar Schools of King Edward VI in Birmingham web site http://www.kingedwardthesixth.org/eligibility.htm)

[23] This information was obtained for 1997 from the Department for Education and Employment website (http://www.dfee.gov.uk/performance/primary_97.htm).

Table 5: Locational variables included in the Hedonic Price Models

Variable	Code	Description and *a priori* Expectations
Proximity to City Centre (secs)	CBD Time	Travel time by car from a property to the city centre. One commonly observed result is that property prices will fall moving away from the city centre.
Proximity to and Size of Local Centres	Local Centre	A weighted average of inverse walking distances to general stores that measure the proximity to local centres whilst accounting for the size of local centres. Expectations are similar to those for proximity to the city centre.
Proximity to a Railway Station (secs)	Railway Station Walk	Walking distance from a property to the nearest railway station. Property values are expected to increase with proximity to a transport network node.
Proximity to a Park (secs)	Park Walk	Walking distance from a property to the nearest park. Property values are expected to increase with proximity to recreational areas.
Proximity to A-Type Industrial Processes (m)	Industry (Part A)	Straight line distance between each property and the nearest large industrial with significant polluting capacity. Since such plants are assumed disamenities it is anticipated that property values will fall with increasing proximity.
Proximity to B-Type Industrial Processes (m)	Industry (Part B)	Straight line distance between each property and the nearest industrial plant with limited pollution capacity. Again, proximity to Part B industrial locations is assumed to be a disamenity, though possible of lesser magnitude to Part A industrial processes.
Proximity to Land Fill sites (m)	Land Fill	Straight line distance between each property and the nearest land fill site. Property values are expected to fall the closer they are to landfills.
Proximity and Quality of Primary Schools	Primary School	Distance weighted average of the performance of nearby primary schools. Since parents may be attracted to locations near better primary schools, it is anticipated that property values will be increasing in this variable.

3.iv. Environmental Characteristics

Road and rail noise data for each property were obtained from the DETR (DETR, 2000). These data were supplied as a noise level on each façade for every residential address in Birmingham that could easily matched to properties in the sample.

The aircraft noise level at each property was identified by digitising a 1999 aircraft noise contour map of Birmingham International Airport. This map displayed day and night time aircraft noise levels in 3dB steps. Each property was assigned a noise level by interpolating linearly between the contours to estimate a noise level at the front of the property.

As described in Table 6, variables for noise pollution from road, rail and aircraft traffic are included in the HPF in a piecewise linear fashion. That is, noise pollution is assumed to have no impact on property prices until it exceeds a threshold level of 55dB. This threshold is often taken as the "background" noise level in urban environments.

Table 6: Environmental variables included in the Hedonic Price Models

Variable	Code	Description and *a priori* Expectations
Road Traffic Noise (dB)	Road Noise	Decibels of noise above 55dB from different sources of noise pollution. The 55dB cut off reflects the fact that noise levels below this level are indistinguishable from "background" noise in an urban environment. Property values are expected to fall in noisier locations.
Railway Traffic Noise (dB)	Rail Noise	
Aircraft Traffic Noise (dB)	Air Noise	
View of Parks	Park View	Distance weighted sum of the area of parkland visible from a property. This is assumed to be an amenity that inflates property prices.
View of Water Surface	Water View	Distance weighted sum of the area of water surface visible from a property. Assumed to be an amenity that inflates property prices.

In order to determine the visual impact of land uses on property prices, GIS were used to construct "viewsheds" for each property. These showed the area of land that was visible from each property by taking into account the height of the land and the heights of features such as buildings. These were then combined with a land use map to determine the area of a variety of land uses visible from each property. Measures of impact were derived by weighting each visible cell by its distance from the observer such that cells nearer to a given property are accorded greater weight in the assessment of visual impact. Details of this procedure can be found in Lake et al. (2000). Variables measuring the visual impact of roads, railways, parks, industry, water, and buildings were calculated from the front of each sample property. We choose to include measures of water views and park views as these amenities are not captured by other variables in the dataset.

4. Factor Analysis of Neighbourhood Characteristics

As illustrated by the numerous attributes listed in Table 3, census data provides a rich source of information on neighbourhood attributes. Indeed, here the problem for hedonic analysis is not one of lack of data on attributes but one of over-abundance. Not surprisingly, many of the attributes listed in Table 3 are highly collinear. For example, the percentage of households not owning cars exhibits high positive correlation with the percentage of households that do not own their own property (correlation coefficient of .86) and high negative correlation with the percentage of households that own two cars (correlation coefficient of -.88).

Whilst each of these neighbourhood attributes might have a bearing on property prices, the presence of such collinearity confounds the estimation of the HPF. Moreover, it is not clear that each of these neighbourhood attributes will be independently capitalised into the property market. More likely, households in a market will consider more general indications of the neighbourhood of a property, the wealth of the area, its ethnic makeup, the stage of life of its inhabitants etc.

As a result, we propose condensing the excess of neighbourhood attributes into a more manageable set of indices. Each index picks out a major dimension of difference or similarity between property neighbourhoods. For example one index might indicate the wealth of a neighbourhood, effectively combining the myriad attributes that are indicators of wealth/poverty into one dimension. Subsequently, property neighbourhoods can be scored along each dimension. In our example, poor neighbourhoods would generate low scores on the wealth dimension, whilst affluent neighbourhoods would generate high scores. The procedure by which dimensions are identified and property neighbourhoods are scored along these dimensions is known as *factor analysis*.

We do not intend presenting the intricacies of factor analysis here (standard texts include Lindeman et al.; 1980). In essence, the procedure seeks to identify major dimensions of association between variables such that a smaller set of variables can be defined that approximate the variation shown in the original data.

Since we are interested in patterns of association, it is not surprising that the first step in a factor analysis is to calculate the correlation matrix of the M variables under study. Each row (or column) of this matrix can be thought of as representing a point in M-dimensional space. We use M axes to locate the points in this space, where each axis represents the degree of correlation with one of the M variables and ranges from -1 to 1. Thus the position of the m^{th} point indicates the nature of the correlation between attribute m and all the other attributes.

If there were no correlation between the attributes, each of the M points would be located on its own axis. Alternatively, when the attributes are correlated, as is the case here, the rows of the correlation matrix form a cloud of points in the space $(-1, 1)^M$. Two attributes showing strong positive correlation will have points located close to each other in this M-space. Likewise, an attribute showing strong negative correlation with these attributes will have a point located near to the mirror image of their points on the opposite side of the origin. Thus, attributes that measure slightly different aspects of one underlying dimension will have points that tend to align themselves along an axis running through the origin.

The first step in a factor analysis is to define an alternative set of axes through this space that capture these patterns of alignment. This is achieved by decomposing the correlation matrix into its eigenvalues and associated eigenvectors. As is well known, the M eigenvectors represent just such a set of alternative orthogonal axes.

Let us consider the eigenvector with the highest associated eigenvalue. It transpires that of all the possible axis that could be drawn through the space $(-1, 1)^M$, the axis defined by the first eigenvector picks out the dimension capturing the most variability in the location of the points. That is, of all possible axes, the first eigenvector distinguishes the most significant alignment of points.

Of course it is unlikely that this one dimension will account for all of the variability in the data. Indeed, the eigenvector with the second highest eigenvalue defines a second axis, orthogonal to the first, best approximating the remaining variation in the points. In the same manner, eigenvectors with successively smaller eigenvalues can be used to better and better approximate the location of the M points. The eigenvectors define the "factors" of factor analysis.

The City of Birmingham has some 1,940 different enumerator districts (EDs) and for each ED our dataset contained details of some 18 neighbourhood attributes (as listed in Table 3). Table 7 details the first eight factors for the ED neighbourhood attributes The second column in this Table provides the eigenvalue of each factor. The third column indicates the percentage of the variation in the location of the M points explained exclusively by that factor. If the attributes were not correlated then each factor would correspond to an original axis and explain $1/M^{th}$ of the variation. If the attributes were all perfectly correlated then they would all be aligned along one axis and this axis would explain 100% of the variation. The fourth column provides the cumulative sum of this explained variation.

Table 7: Variation explained by the first ten factors of the Enumerator District neighbourhood attributes (estimated using Iterated Principal Factors)

Factor	Eigenvalue	Variation Explained by Factor	Cumulative Explained Variation
1	7.28	.43	.44
2	3.24	.20	.63
3	1.79	.11	.74
4	1.21	.25	.81
5	0.95	.30	.87
6	0.65	.11	.91
7	0.55	.15	.94
8	0.39	.17	.97

From Table 7 it can be seen that the first factor alone explains over 40% of the variation in the ED data. This indicates that many of the attributes are highly correlated with a single underlying factor. Notice that successive factors explain progressively less of the remaining variation. A good rule of thumb for selecting factors, is to choose those factors with eigenvalues greater than one. This procedure leads us to focus on the first 4 factors. As such, our constructed indices will explain around 81% of the variation in the neighbourhood attribute data.

One of the arts of factor analysis is the interpretation of factors. Interpretation of factors is the process of describing the underlying dimension of similarity or difference between the neighbourhood attributes captured by a factor. To explain how this is achieved, observe that if the axis defined by a particular factor is closely aligned with an original attribute axis, then that attribute is important in determining the factor. Conversely if the factor axis is orthogonal to an original axis, the attribute described by that axis plays no part in determining that factor. The degree to which individual attributes contribute to a factor is measured by the factor loadings.[24] A large positive loading indicates that high values of the original attribute are associated with high values of the factor. Similarly a large negative loading indicates that high values of the original attribute are associated with low values of the factor. The factor loadings for our application are listed in Table 8.

Table 8: Rotated Factor Loadings

Attribute	Factor 1	Factor 2	Factor 3	Factor 4
No car	-0.96	+0.15	-0.07	+0.09
Two cars	+0.85	-0.21	+0.09	+0.02
Unemployment	-0.79	+0.35	-0.13	+0.19
Non-owners	-0.89	-0.08	-0.03	+0.05
One-parent families	-0.72	-0.10	-0.27	+0.38
Low Social Class	-0.44	+0.17	-0.02	+0.09
Families	-0.10	+0.38	+0.12	+0.80
Age 0 to 10	-0.39	+0.30	-0.33	+0.74
Age 11 to 17	+0.18	+0.54	+0.16	+0.68
Age 18 to 24	-0.30	+0.33	-0.62	-0.16
Age 25 to 34	-0.15	-0.06	-0.85	-0.09
Age 35 to 49	+0.80	-0.30	+0.00	+0.11
Age 50 to 64	+0.27	-0.04	+0.57	-0.43
Age > 65	-0.22	-0.35	+0.64	-0.48
Over Crowding	-0.30	+0.76	-0.04	+0.37
Non White	-0.23	+0.92	-0.14	+0.19
Black	-0.48	+0.40	-0.29	+0.10
Asian	-0.08	+0.94	-0.04	+0.20

The loadings in Table 8 suggest fairly obvious interpretations for the four factors.

[24] Frequently researchers will *rotate* the factor axes to improve the ease with which factors can be interpreted. That is, the subspace defined by the factor axes is not changed, but the orientation of the axes themselves are rotated such that they best align with original axes describing the attributes.

We suggest the following;

Factor 1: Wealth

This factor is very distinct and describes the general level of wealth of neighbourhoods. Not surprisingly, the factor is highly positively correlated with car ownership and being in the maximum earning bracket age range between 35 and 49. Conversely the factor is strongly negatively associated with lack of access to cars, unemployment, low home ownership and one parent families.

Factor 2: Ethnicity

The second factor loads heavily on four attributes; the three attributes describing the ethnic composition of neighbourhoods and the attribute describing the degree of over-crowding in households. Since all four of these loadings are positively signed, high scores on this dimension reflect the increasing presence of members of the ethnic minorities in neighbourhoods.

Factor 3: Adult Age Composition

This fourth factor picks out a dimension defining the age composition of neighbourhoods. It loads negatively on young adults but loads positively on adults in older generations. EDs scoring highly on this factor will be characterised by neighbourhoods with relatively older adult populations.

Factor 4: Family Composition

The final factor loads heavily on just three attributes those describing the percentage of households with children and the percentage of the ED populations in age groups 0 to 10 and 11 to 17. EDs that score highly on this factor are characterised by having a relatively large number of households with children. Notice the distinction here with the composition of adult ages as described by the third factor. Clearly, it is possible to have EDs exhibiting the same distribution of adult age ranges but which differ according to the degree to which those adults have children.

The final step in a factor analysis is to use the factor loadings to define a score for each ED for each factor. Using the factor loadings a regression-like equation is calculated, the parameters of which indicate how greatly each attribute contributes to each factor. Given the attributes of each neighbourhood, the equation can be used to determine how highly a neighbourhood scores on each factor. In effect, neighbourhoods that exhibit high values for attributes that load positively on a factor receive high scores for that factor whilst neighbourhoods that exhibit high values for attributes that load negatively on that factor receive low scores. Details of the factor scores are provided in Table B1 in Appendix B.

As detailed in Table 9, the factor scores are used as proxies for the original attributes in regression analysis. Further, in Section 6 of this Chapter we employ the factors as indicators of the socioeconomic characteristics of property neighbourhoods in order to identify market segments. As has been demonstrated the factors capture a good proportion of the variation shown in the original neighbourhood attributes. Moreover, the nature of their construction ensures that the factor scores are orthogonal thereby overcoming problems induced by collinearity in the original set of attributes.

Table 9: Neighbourhood variables included in the Hedonic Price Models

Variable	Code	Description and *a priori* Expectations
Wealth Factor	Wealth	Higher scores indicate wealthier neighbourhoods. Properties in wealthier neighbourhoods are expected to command higher prices.
Ethnicity Factor	Ethnicity	Higher scores indicate neighbourhoods with a higher percentage of households from the ethnic minorities. If households prefer to locate in neighbourhoods whose residents are culturally similar to themselves then the influence of ethnicity on property prices will be determined by the ethnic make up of each market segment.
Age Composition Factor	Age Composition	Higher scores indicate neighbourhoods with older adult residents. Though the influence on age composition is uncertain, it is expected that properties in neighbourhoods with generally older residents will command higher prices.
Family Composition Factor	Family Composition	Higher scores indicate neighbourhoods with a higher proportion of households with children. Though the influence on age composition is uncertain, it is expected that properties in neighbourhoods with more families are expected to command lower prices.

5. Summary of Data

The City of Birmingham property dataset is perhaps the richest of its kind yet to be constructed for any property market. Actual sale prices have been provided by the UK Land Registry and Geographical Information Systems (GIS) techniques used to locate these properties on digital maps.

Descriptions of the variables used in the hedonic analysis are listed in Table 10. Complete data records were successfully compiled for some 10,848 residential property transactions in Birmingham in 1997.

Table 10: Data Descriptions

Variable	Mean	Std. Dev.	Min	Max
Sale Price (£)	58,986	36,099	11,000	645,003
Structural Characteristics				
Floor Area (m^2)	102.6	32.7	42	645
Garden Area (m^2)	226.1	208	0	5,164
Garage (proportion)	0.436	0.496	0	1
Central Heating (proportion)	0.728	0.268	0	1
Age (decades)	6.1	2.76	0	11
WCs (proportion)				
One	0.794	0.404	0	1
Two	0.196	0.397	0	1
Three	0.009	0.094	0	1
> Three	0.001	0.029	0	1
Bedrooms (proportion)				
One	0.005	0.069	0	1
Two	0.172	0.377	0	1
Three	0.716	0.451	0	1
Four	0.083	0.276	0	1
Five	0.016	0.127	0	1
> Five	0.007	0.084	0	1
Storeys (proportion)				
One	0.021	0.145	0	1
Two	0.954	0.209	0	1
Three	0.021	0.143	0	1
> Three	0.003	0.058	0	1

Variable	Mean	Std. Dev.	Min	Max
Construction Type (proportion)				
Detached Bungalow	0.013	0.111	0	1
Semi-Detached Bungalow	0.008	0.090	0	1
End Terrace Bungalow	0.000	0.022	0	1
Terrace Bungalow	0.000	0.017	0	1
Detached House	0.116	0.320	0	1
Semi-Detached House	0.396	0.489	0	1
End Terrace House	0.115	0.319	0	1
Terrace House	0.352	0.478	0	1
Beacon Group (proportion)				
1. Unrenovated cottage pre 1919	0.000	0.019	0	1
2. Renovated cottage pre 1919	0.001	0.027	0	1
3. Small "industrial" pre 1919	0.040	0.195	0	1
4. Medium "industrial" pre 1919	0.226	0.418	0	1
5. Large terrace pre 1919	0.006	0.078	0	1
8. Small "villa" pre 1919	0.020	0.138	0	1
9. Large "villas" pre 1919	0.009	0.093	0	1
10. Large detached pre 1919	0.003	0.058	0	1
19. Houses 1908 to 1930	0.011	0.103	0	1
20. Subsidy houses 1920s & 30s	0.140	0.347	0	1
21. Standard houses 1919-45	0.257	0.437	0	1
24. Large houses 1919-45	0.016	0.124	0	1
25. Individual houses 1919-45	0.000	0.022	0	1
30. Standard houses 1945-53	0.045	0.207	0	1
31. Standard houses post 1953	0.190	0.392	0	1
32. Large houses post 1953	0.032	0.177	0	1
35. Individual houses post 1945	0.001	0.038	0	1
36. "Town Houses" post 1950	0.004	0.062	0	1
Sale Date (proportion)				
1st Quarter (Jan. to Mar.)	0.214	0.410	0	1
2nd Quarter (Apr. to June)	0.247	0.431	0	1
3rd Quarter (July to Sept.)	0.287	0.452	0	1
4th Quarter (Oct. to Dec.)	0.252	0.434	0	1
Neighbourhood Characteristics				
Poverty Factor	-0.375	0.855	-1.934	2.363
Sills Factor	0.180	1.000	-1.398	4.198

Variable	Mean	Std. Dev.	Min	Max
Age Factor	0.055	0.807	-3.216	3.143
Family Factor	-0.029	0.842	-3.198	3.791
Asian Factor	-0.045	0.942	-1.131	5.152
Black Factor	-0.240	0.750	-2.016	8.214
Locational Characteristics				
Proximity to City Centre (mins)	1,313	478	208	3,187
Proximity and Quantity of Shops	2.276	1.273	0.07	9.56
Proximity and Quality of Primary Schools	0.602	0.177	0.15	0.97
Walking time to Rail Station (secs)	1,846	1,013	21.05	5,525
Walking time to a Park (secs)	900	558	3.17	3,425
Driving time to Airport (secs)	2,388.215	655.134	602.19	4,386
Proximity to A-Type Industrial Processes (m)	2,463.592	1,820.591	21.94	10,204
Proximity to B-Type Industrial Processes (m)	814.103	527.842	10	3,333
Proximity to Land Fill sites (m)	946.611	608.089	10	3,472
Wards (proportion)				
Acock's Green	0.039	0.194	0	1
Aston	0.015	0.122	0	1
Bartley Green	0.018	0.131	0	1
Billesley	0.027	0.162	0	1
Bournville	0.038	0.191	0	1
Brandwood	0.022	0.147	0	1
Edgbaston	0.020	0.139	0	1
Erdington	0.029	0.168	0	1
Fox Hollies	0.028	0.165	0	1
Hall Green	0.041	0.198	0	1
Handsworth	0.016	0.125	0	1
Harborne	0.036	0.186	0	1
Hodge Hill	0.024	0.154	0	1
King's Norton	0.016	0.125	0	1
Kingsbury	0.010	0.101	0	1
Kingstanding	0.022	0.146	0	1
Ladywood	0.014	0.118	0	1
Longbridge	0.023	0.150	0	1
Moseley	0.024	0.152	0	1

Variable	Mean	Std. Dev.	Min	Max
Nechells	0.019	0.136	0	1
Northfield	0.028	0.164	0	1
Oscott	0.026	0.158	0	1
Perry Barr	0.033	0.180	0	1
Quinton	0.024	0.152	0	1
Sandwell	0.027	0.163	0	1
Selly Oak	0.044	0.205	0	1
Shard End	0.020	0.138	0	1
Sheldon	0.021	0.144	0	1
Small Heath	0.028	0.164	0	1
Soho	0.018	0.134	0	1
Sparkbrook	0.013	0.111	0	1
Sparkhill	0.020	0.142	0	1
Stockland Green	0.028	0.166	0	1
Sutton Four Oaks	0.038	0.191	0	1
Sutton New Hall	0.044	0.206	0	1
Sutton Vesey	0.039	0.194	0	1
Washwood Heath	0.028	0.164	0	1
Weoley	0.017	0.130	0	1
Yardley	0.024	0.154	0	1
Environmental Characteristics				
Views of Water (weighted m^2)	0.480	7.543	0	348
Views of Parkland (weighted m^2)	6.290	36.831	0	664
Road Traffic Noise (dB)	49.8	9.4	31.6	75.8
Rail Traffic Noise (dB)	36.8	12.6	0	74.7
Aircraft Noise (dB)	4.8	16.0	0	69

6. Identifying Market Segments using Cluster Analysis

6.i. Introduction

There are a number of theoretical and empirical reasons why accounting for market segmentation is both a desirable and necessary step in hedonic analysis. Rather than repeating the discussion in Section 3 of Chapter 3, we note that if more than one market segment exists in the City of Birmingham property market then it is likely that the HPFs for each segment are quite different. Separate hedonic functions should be estimated for each segment as failing to differentiate between segments may seriously bias the parameter estimates.

Whilst identifying a separate price functions for each market segment complicates the analysis, it also brings benefits. In particular, different households within Birmingham find themselves purchasing properties in different property markets boasting different HPFs. As we shall describe in Part 3 of this thesis, identification of a demand curve for property characteristics such as noise pollution, will only be possible in the presence of data drawn from several such market segments.

A number of researchers have investigated the existence of market segments in urban areas (e.g. Straszheim, 1973; Schnare and Struyk, 1976; Ball and Kirwan, 1977; Sonstelie and Portney, 1980; Goodman, 1978; Micheals and Smith, 1990; Allen et al., 1995). These studies have applied different rules by which properties in an urban area are allotted to a particular market segment. Criteria include; locational or political boundaries, characteristics of households (e.g income and race), property types and classifications based upon the judgement of estate agents.

However, as pointed out by Watkins (2001), it is unlikely that any one criterion is sufficient to characterize market segments. Alternatively properties could be grouped into clusters based on their similarity along a multitude of dimensions; locational, structural and socioeconomic. One data-driven process by which such groups might be identified is known as *cluster analysis*. Indeed techniques of cluster analysis have seen some application to the classification of properties into market segments; notably Abraham et al. (1994), Goetzmann and Wachter (1995), Hoesli et al. (1997), Bourassa et al. (1999), Day (2003) and Day et al. (2003). However, these studies all use relatively simple clustering algorithms that provide no independent statistical indication of the nature or number of clusters to be found in the data. In this research we take advantage of recent advances in clustering techniques to answer both these important questions. In particular, we pioneer the use of *model-based cluster analysis* and a description and application of this technique forms the subject matter of this Section.

6.ii. Cluster Analysis

Simply stated, the objective of cluster analysis is to group together observations that are similar to each other (with respect to a set of P variables) into clusters. Of course, there are as many approaches to cluster analysis as there are definitions of "similarity" and the constitution of a "cluster".

A common starting point is to define each observation as a point in P-space whose location is determined by how highly that observation scores for each variable. Clearly, observations holding similar values for the different variables will be located close to each other in this P-space. The similarity of two observations can then be judged by a metric such as the Euclidian or Mahalanobis distance between the observations.

Until recently clustering approaches tend to have been based around intuitively simple iterative algorithms (details may be found in Jain and Dubes, 1988, or Everitt, 1993). For example, hierarchical clusters are often constructed using simple algorithms in which, at each stage of the cluster process, the two nearest clusters are merged.[25] Alternatively, clusters can be constructed using partitioning algorithms. Here observations are initially divided into a given number of clusters and then iteratively reallocated (i.e. moved between clusters) until a desired goal is reached.[26]

Unfortunately, relatively little is known concerning the various statistical properties of this collection of clustering algorithms. As a result, little can be said concerning the relative merits of different clustering solutions, neither in terms of the number of clusters identified in the data nor in terms of the shape, size and composition of the identified clusters.

In recent years a number of new approaches to identifying clusters in data have been proposed (reviewed in Fasulo, 1999, and Jain et al., 1999). One approach that has shown particular promise is that of model-based clustering (McLachan and Basford,

---

[25] For example, Ward (1963) proposed that the two clusters are merged that result in the smallest increase in the sum of squared distances between each point and the cluster centre.

[26] The classic k-means algorithm (MacQueen, 1967; Hartigan, 1975; Hartigan and Wong, 1978), for example, takes a fixed number of k clusters and assigns observations to those clusters so that the centroids of the clusters are as different from each other as possible.

1988; Banfield and Raftery, 1993; Fraley and Raftery, 1998; Fraley and Raftery, 2002a). This clustering approach has been successfully applied to a variety of data problems across a broad range of disciplines. For example, in the biological sciences to analyse gene expression data (e.g. Ghosh and Chinnaiyan, 2002; Lin et al., 2002; McLachlan et al., 2002; Yeung et al., 2001), in ecology to study community composition (e.g. Ter Braak et al., 2003), in atmospheric sciences to study circulation patterns (e.g. Smyth, 2000; Smyth et al., 1999), in astronomy to classify gamma ray bursts (e.g. Mukherjee et al, 1998) and in various fields for image analysis (e.g. Gopal and Hebet, 1998; Campbell et al, 1999; Wehrens et al., 2002). In contrast, model-based clustering techniques are relatively unknown in economic analysis. As far as the author is aware this is the first application of these techniques in this field.

6.iii. Model-Based Cluster Analysis

The fundamental assumption of model-based clustering is that each data point is drawn from a population of such points constituting all the members of the cluster. Moreover, the location, size and shape of this underlying population can be approximated by a probability distribution. Assuming a Gaussian distribution, for example, would imply that clusters are ellipsoidal. It would also assume that the likelihood of observing data points belonging to a particular cluster is greater near the mean location of that cluster than at its periphery. The data observed by the researcher is the composite of data points drawn from a finite number of such clusters.

To formalise, each P-dimensional data point $\boldsymbol{x}$ arises from a super population comprising a mixture of M populations, $C_1, C_2, \ldots, C_M$, in some proportions $\pi_1, \pi_2, \ldots, \pi_M$ respectively, where;

$$\sum_{j=1}^{M} \pi_j = 1 \quad \text{and} \quad \pi_j \geq 0 \quad (j = 1, 2, \ldots, M) \tag{3}$$

If we assume that each population, C_j, can be modelled as a p-dimensional Gaussian distribution with mean vector $\boldsymbol{\mu}_j$ and covariance matrix $\boldsymbol{\Sigma}_j$, then the probability density function (pdf) of an observation $\boldsymbol{x}$ is of the form;

$$
\begin{aligned}
f(\boldsymbol{x} \mid \boldsymbol{\theta},\boldsymbol{\pi}) &= \sum_{j=1}^{M} \pi_j f_j(\boldsymbol{x} \mid \boldsymbol{\theta}_j) = \sum_{j=1}^{M} \pi_j \phi_j^p(\boldsymbol{x} \mid \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j) \\
&= \sum_{j=1}^{M} \pi_j \left(2\pi^p \left|\boldsymbol{\Sigma}_j\right|\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\left(\boldsymbol{x}-\boldsymbol{\mu}_j\right)' \boldsymbol{\Sigma}_j^{-1}\left(\boldsymbol{x}-\boldsymbol{\mu}_j\right)\right)
\end{aligned} \tag{4}
$$

where $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_M)$ is the vector of parameters associated with the assumed distributions of the M clusters, $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_M)$ is the vector of mixing proportions, $f_j(\boldsymbol{x} \mid \boldsymbol{\theta}_j)$ is the pdf of cluster C_j which is given the specific p-dimensional Gaussian form denoted $\phi_j^p(\boldsymbol{x} \mid \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$. Thus, in the Gaussian case $\boldsymbol{\theta}_j$ comprises the elements of the vector $\boldsymbol{\mu}_j$, which determine the mean location of each cluster in p-space, and the distinct elements of the covariance matrices $\boldsymbol{\Sigma}_j$, which determine the geometric proportions of each cluster.

To allow for comparison of different assumptions concerning the geometric characteristics of the different clusters, Banfield and Raferty (1993) reparameterise each covariance matrix $\boldsymbol{\Sigma}_j$ using the eigenvalue decomposition;

$$
\boldsymbol{\Sigma}_j = \lambda_j \boldsymbol{D}_j \boldsymbol{A}_j \boldsymbol{D}_j' \quad (j = 1, 2, \ldots, M) \tag{5}
$$

where $\boldsymbol{D}_j$ is the matrix of eigenvectors, λ_j is the first eigenvalue of $\boldsymbol{\Sigma}_j$, and $\boldsymbol{A}_j$ is a diagonal matrix with diagonal elements $1 = \alpha_{1j} \geq \alpha_{2j} \geq \ldots \geq \alpha_{pj} > 0$.

The advantage of Banfield and Raferty's decomposition is to isolate different geometric properties of each cluster into different components. Hence λ_j determines cluster volume, $\boldsymbol{D}_j$ cluster orientation and $\boldsymbol{A}_j$ other properties of the cluster shape. Thus imposing the restriction $\lambda_j = \lambda$ $(j = 1, 2, \ldots, M)$ enforces equality of volume across all clusters. Similarly, imposing the restriction $A_j = I$ $(j = 1, 2, \ldots, M)$, where I is the p-dimensional identity matrix, generates strictly spherical clusters. Clearly differing combinations of restrictions imply different imposed similarities between clusters. As we shall see shortly, the great advantage of model-based clustering is that it provides a formal framework in which such restrictions can be compared.

For now, imagine that the number of clusters, M, in the data is known. Also imagine we know to which of these clusters each data point belongs. In that case the complete-data log likelihood can be written;

$$\ln L(\boldsymbol{\theta},\boldsymbol{\pi}) = \sum_{i=1}^{N}\sum_{j=1}^{M} d_{ij} \ln\left[\pi_j \phi_j^p\left(\boldsymbol{x} \mid \boldsymbol{\theta}_j\right)\right] \tag{6}$$

where $\boldsymbol{\theta}_j = (\boldsymbol{\mu}_j,\lambda_j,\boldsymbol{A}_j,\boldsymbol{D}_j)$ $(j = 1, 2, \ldots, M)$ and d_{ij} $(i = 1, 2, \ldots, N; j = 1, 2, \ldots, M)$
where $\boldsymbol{\theta}_j = \left(\boldsymbol{\mu}_j,\lambda_j,\boldsymbol{A}_j,\boldsymbol{D}_j\right)$ $\left(j = 1, 2, \ldots, M\right)$ and d_{ij} $\left(i = 1, 2, \ldots, N; j = 1, 2, \ldots, M\right)$ are indicator variables whose value is 1 if observation i belongs to cluster C_j and 0 otherwise.

Of course, we do not know the provenance of each data point; from the researcher's point of view the d_{ij} are missing data. As Celeux and Govaert (1995) describe, this motivates a simple application of the EM algorithm (Dempster et al., 1977).

The E-step of the algorithm calculates;

$$\bar{d}_{im} = E\left[d_{im} \mid \boldsymbol{x}_i,\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\pi}}\right] = \frac{\hat{\pi}_m \phi\left(\boldsymbol{x}_i \mid \hat{\boldsymbol{\theta}}_m\right)}{\sum_{j=1}^{M}\hat{\pi}_j \phi\left(\boldsymbol{x}_i \mid \hat{\boldsymbol{\theta}}_j\right)} \quad \left(m = 1, 2, \ldots, M\right) \tag{7}$$

where $\bar{d}_{im}$ is the expected value of the indicator variable for membership of cluster m conditional on the data and current parameter estimates denoted $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\pi}}$ with cluster specific components $\hat{\boldsymbol{\theta}}_j$ and $\hat{\pi}_j$ respectively.

In the M-step, $\bar{d}_{im}$ replaces d_{im} in the complete-data log-likelihood (6), which is then maximised with respect to the parameters. Solution of this maximisation problem provides simple closed forms for the mean cluster locations and mixing probabilities;

$$\hat{\boldsymbol{\mu}}_m = \frac{\sum_{i=1}^{N}\bar{d}_{im}\boldsymbol{x}_i}{N_m}; \quad \hat{\pi}_m = \frac{N_m}{N}; \quad N_m \equiv \sum_{i=1}^{N}\bar{d}_{im} \quad \left(m = 1, 2, \ldots, M\right) \tag{8}$$

Estimating the elements of the covariance matrix $\hat{\Sigma}_m$, in the M-Step, depends on the particular parameterisation. Further details of these computations using the eigenvalue decomposition in (5) can be found in Celeux and Govaert (1995). The E-step and M-step are iterated until convergence of the parameters.

The value $\bar{d}_{im}^*$ of $\bar{d}_{im}$ that maximises (6) gives the conditional probability that observation i belongs to cluster C_m. A maximum likelihood classification of the data can be derived by associating each observation with the cluster to which it is most likely to belong. That is, observation i is classified as belonging to cluster C_m if $d_{im}^* = \max_j d_{ij}^*$. Furthermore, $1 - \max_j d_{ij}^*$ gives a measure of the uncertainty associated with each observation's classification (Bensmail et al., 1997)

As is clear from (7) and (8) the EM algorithm decomposes the problem of maximising the mixture model log-likelihood (6) into a series of relatively simple calculations. As described by Fraley and Raftery (2001), this simplicity comes at a cost. In particular, the conditions under which the algorithm can be proven to converge to a local maximum do not always hold for mixture models. Nonetheless, Fraley and Raftery (2001) indicate that EM estimation has been applied with considerable success in this context. Furthermore, the rate of convergence of the algorithm may be very slow and may encounter difficulties if there are a large number of clusters or the data is ill-conditioned. As with all maximisation problems, the chances of reaching a satisfactory solution are greatly enhanced by initialising the algorithm with reasonable starting values, a subject we shall return to discuss shortly.

6.iv. Model selection

One problem that remains is how to choose between clustering solutions allowing different numbers of components and differing parameterisations of cluster shapes. In contrast to other clustering algorithms, the probabilistic basis of model-based clustering provides a framework within which these comparisons can be made.

A Bayesian approach to model selection is to choose the model that is most likely *a posteriori*. Given that *a priori* all models are considered equally likely, this amounts to comparing the integrated likelihood of the different models. Unfortunately, even for relatively simple Gaussian mixture models this integral has no closed form. An alternative then, is to use penalized likelihood methods that approximate the integrated likelihood. One such method is the Bayesian Information Criterion (BIC) (Schwartz, 1978);

$$BIC_g = 2 \ln L\left(\boldsymbol{\pi}_g^*, \boldsymbol{\theta}_g^*\right) - \upsilon_g \log(N) \tag{9}$$

where g indexes the particular model being evaluated, $\boldsymbol{\pi}_g^*$ and $\boldsymbol{\theta}_g^*$ are the maximum likelihood estimates of $\boldsymbol{\pi}$ and $\boldsymbol{\theta}$ respectively and υ_g are the total number of independent parameters in $\boldsymbol{\pi}_g^*$ and $\boldsymbol{\theta}_g^*$. If a BIC statistic is calculated for two different models, the difference between their BICs is what will indicate the superiority of one model over the other. If the difference is large enough, one can be reasonably certain that one model gives a better fit than the other. A standard convention for calibrating BIC differences is that differences of less than 2 correspond to weak evidence, differences between 2 and 6 to positive evidence, differences between 6 and 10 to strong evidence, and differences greater than 10 to very strong evidence.

Whilst, the regularity conditions necessary for the BIC to approximate the integrated likelihood do not hold for finite mixture models (Titterginton et al., 1986), a growing weight of theoretical and empirical evidence supports the use of the BIC in this context (Leroux, 1992; Keribin, 1998; Roeder and Wasserman, 1997; Campbell et al., 1999; Dasgupta and Raftery, 1998; Fraley and Raftery, 1998, 2002; Stanford and Raftery, 2000).[27]

The approach followed here is that outlined in Fraley and Raftery (1998). In the first instance select a range for M, the number of clusters. Then select a series of parameterizations of the covariance matrix by applying one or more equality restrictions to (5). For each value of M and each parameterization, use the EM

[27] Other approaches to model comparison include the NEC an entropy criterion proposed by Biernacki et al. (1999), a bootstrap approach followed by McClachlan (1987) and the cross-validated likelihood approach of Symthe (2000). Comparisons of some of these different approaches can be found in Pan et al. (2002) and Biernacki and Govaert (2000).

algorithm to calculate the maximum likelihood estimates of the model parameters. Compute the BIC for each model. The model providing the highest value of the BIC is then selected.

In large datasets, the BIC tends to favour models with many clusters (Posse, 2001). Thus we follow the example of Posse (2001) who suggests "picking a good candidate in the region where the rate of change of the BIC drops significantly".

6.v. Initialisation of the EM Algorithm

Since the likelihood surface is typically characterised by many local maxima, finding appropriate starting values for the EM algorithm is a very important issue (Biernacki et al., 2003, compare various initialisation strategies). One approach is to derive starting values from an initial hierarchical clustering of the data (Fraley and Raftery, 1998). In particular, Fraley and Raftery (1998) propose initialisation using model-based hierarchical clustering (Banfield and Raftery, 1993) with an unconstrained covariance matrix. Here each observation begins in a cluster of its own and, at each stage, a pair of clusters are merged so as to maximise a log-likelihood function (see Banfield and Raftery, 1993, for details). Each step in the hierarchical clustering defines a unique number of clusters until in the final step all observations are tied together in one cluster. The output from this hierarchical clustering can be illustrated as a dendrogram revealing the association between observations.

To categorise the observations into M partitions, a section can be taken through the dendrogram at the level isolating M clusters. Fraley and Raftery (1998) propose using this categorisation to provide starting values for the cluster membership indicators d_{ij} $(i = 1, 2, \ldots, N; j = 1, 2, \ldots, M)$. These, in turn, can replace the conditional probabilities (7) to feed into the initial M-Step of the EM iteration.

One shortcoming of this approach is the onerous computing requirements of hierarchical clustering methods. In particular, the initial step of agglomerative hierarchical methods requires a measure of distance to be calculated between each observation in the data. As a result computing time and storage requirements are at least quadratic in the number of observations. Indeed, hierarchical clustering of large datasets may prove unfeasible.

Two basic approaches have been forwarded to overcome these constraints. Banfield and Raftery (1993) propose clustering a subsample of the data then using discriminant analysis to classify the remaining observations (see also Maitra, 2001; Buydens and Raftery, 2003). Alternatively, Posse (2001) suggests an approach that takes into account all of the observations in the dataset. Rather than beginning the hierarchical agglomeration from the set of singleton clusters, Posse (2001) proposes initially categorising observations into a smaller number of well-defined clusters. Provided the initial clustering is efficient, such that it only groups observations that would naturally fall into the same cluster at a relatively early stage in the hierarchical agglomeration, this approach should result in a similar classification as that achieved through a hierarchical clustering of the entire dataset.

Here we follow the Posse (2001) approach that draws on graph theoretic approaches to clustering. In particular, Posse (2001) suggests generating clusters from the minimum spanning tree (MST) of the data. A spanning tree is a graph that connects all the data points in P-space such that there is only one path connecting each pair of data points. The MST is the spanning tree in which the total length of the connections or edges joining each point is at a minimum.

Posse's (2001) approach involves two steps in which the MST is first "peeled" and then "pruned". The peeling step involves trimming out the longest edges of the MST. In effect this divides the well-separated groups in the dataset into discrete clusters whilst also isolating observations on the periphery of clusters that would not be assigned to a cluster until late on in the hierarchical agglomeration. The pruning step involves dividing the surviving connected observations in the MST into small groups each of roughly the same size. These groups should consist of close neighbours that would have been merged early on in the hierarchical clustering. Observations that are connected after peeling and pruning are given the same classification and this acts as the partition from which the hierarchical clustering is initiated.

To determine which edges in the MST are considered sufficiently long to warrant "peeling", Posse (2001) proposes the use of plots comparing the observed distribution of the longest edge lengths in the data with those that would be expected if the data had come from a single Gaussian population. To this end, Posse (2001) extends a theorem of Penrose (1998) that describes the expected distribution of edge lengths in the MST if data points were drawn from a single standard p-dimensional

Gaussian distribution. In particular, the theorem states that for the standard Gaussian distribution, the probability of observing the l^{th} longest edge in the MST to have length e_l, is given by the l^{th} order Gumbel distribution once e_l has been suitably centred and scaled. More formally;

$$\lim_{N\to\infty} P\left[a_N e_l - b_N \le x\right] = G_l(x)$$

where

$$\begin{aligned}
G_l(x) &= \exp\left(-e^{x}\right)\left(1+\sum\nolimits_{j=1}^{l-1}\exp(-jx)/\Gamma(l)\right)\\
a_N &= 2^{1/2}\left(\ln N + ((p/2)-1)\ln_2 N \;-\; \ln\Gamma(p/2)\right)^{1/2}\\
b_N &= (p-1)\ln_2 N \;-\; ((p-1)/2)\ln_3 N \;-\; \ln\kappa_p\\
\kappa_p &= 2^{-p/2}(2\pi)^{-1/2}\Gamma(p/2)(p-1)^{(p-1)/2}
\end{aligned}$$

and

$$\ln_2 N = \ln\ln N,\quad \ln_3 N = \ln\ln\ln N \qquad (10)$$

Thus, the quantities $u_1 = G_1(a_N e_1 - b_N), u_2 = G_2(a_N e_2 - b_N), \ldots, u_N = G_N(a_N e_N - b_N)$ are identically (though not independently) distributed according to the uniform distribution in the interval $[0,1]$.[28]

However, if the data are not in fact drawn from a single standardised Gaussian population the u_l sequence will not show this pattern. As Posse (2001) describes, if separate clusters are present in the data, the ordered sequence of edge lengths given by e_l will tend to be longer than expected for early elements in the sequence. Similarly, if clusters are not homogenous but are elongated in some dimension then early values of the edge length sequence e_l will also be longer than expected. As a consequence, in either or both of these cases, the observed u_l sequence will be characterised by initial values close to 1 before decreasing rapidly towards 0.

---

[28] Note that the u_l sequence can be easily calculated since for $l > 7$ the G_l distribution is accurately approximated by the Gaussian distribution with mean $\mu_l = -\ln l + 1/2l$ and standard deviation $\sigma_l - (l-1/2)^{-1/2}$. Further, Posse (2001) notes the slow rate of convergence of the limit in (10) and provides second order corrective terms for a_n and b_n obtained from Monte-Carlo simulations. As acknowledged by the author, there is a small error in Posse's (2001) equation (6) where the Monte Carlo correction term should in fact be subtracted from b_n rather than added.

Posse (2001) suggests that the number of edges to be peeled should be determined by plotting both the u_l sequence and the e_l sequence. These plots should reveal the point at which the u_l sequence stabilises around 0 and the point at which the rate of decay in the e_l sequence drops significantly. Posse (2001) indicates that a suitable choice for the number of edges to peel is the largest of these two quantities.

6.vi. Overall clustering strategy

The clustering strategy followed in this paper, therefore, follows a number of steps;

1. Construct the MST. In our application the MST is constructed using Prim's (1957) algorithm. To account for different scaling in the p clustering variables, inter-point distance is measured using a Mahalanobis metric.
2. Peel and prune the MST. Observations that are still connected in the MST following peeling and pruning represent a good initial partition of the data.
3. Perform a hierarchical clustering of the data. An agglomerative model-based hierarchical clustering is performed on the data, starting from the initial partition determined in step 2
4. Determine the number of clusters, M, and a parameterisation for the cluster covariance matrices.
5. Perform a model-based clustering of the data using the EM algorithm. The EM algorithm is begun at the M-Step with the cluster membership indicators d_{ij} initialised from a classification of observations corresponding to M clusters taken from the hierarchical clustering in step 3.
6. Calculate the BIC for this model
7. Repeat steps 4 to 6 for various numbers of clusters and parameterisations of the cluster covariance matrices.
8. Plot the values of the BIC for the different models and choose a good candidate model as that giving the highest value for the BIC in the region where the rate of change of the BIC drops significantly.

6.vii. Geographic smoothing

One last issue remains to be resolved. We might expect that for properties, geographical location will play an important role in determining market segment membership. Unfortunately directly including locational variables in a cluster analysis when observations are spread reasonably homogeneously across space, tends to result in a large number of clusters that are nearly circular when spatially mapped (Fovell, 1997). As a result we follow Posse (2001) and post-process the clustering classification to take account of the spatial information.

Here we adopt a very simple rule. The eight closest observations in geographical space to each observation are identified. The nine observations are examined and their classification noted. If the majority of these observations favour one classification and this differs from the classification of the target observation then the probabilities of belonging to these two different clusters (as given by the respective values of d^*_{im}) are compared. Only if the target observation is less that twice as likely to belong to its current classification is the classification switched. This spatial smoothing rule is applied to all observations and the process iterated until no observations change classification.

6.viii. Software

Software implementing the MST initial partition has been written by the author in the GAUSS programming language. This code has subsequently been verified through comparison with Christian Posse's original code designed to interface with the S-PLUS software package.

The model-based clustering (both hierarchical and EM) and BIC calculation have been implemented using Fraley and Raftery's (2002) MCLUST package designed to interface with either S-PLUS or R.

The simple spatial smoothing routine was written by the author in the GAUSS programming language.

7. Results of the Cluster Analysis and Market Segment Descriptions

7.i. Choice of clustering variables

As described previously, different researchers have used a variety of criteria by which to classify properties into market segments. In general, these criteria tend to reflect one or more of the following categories;

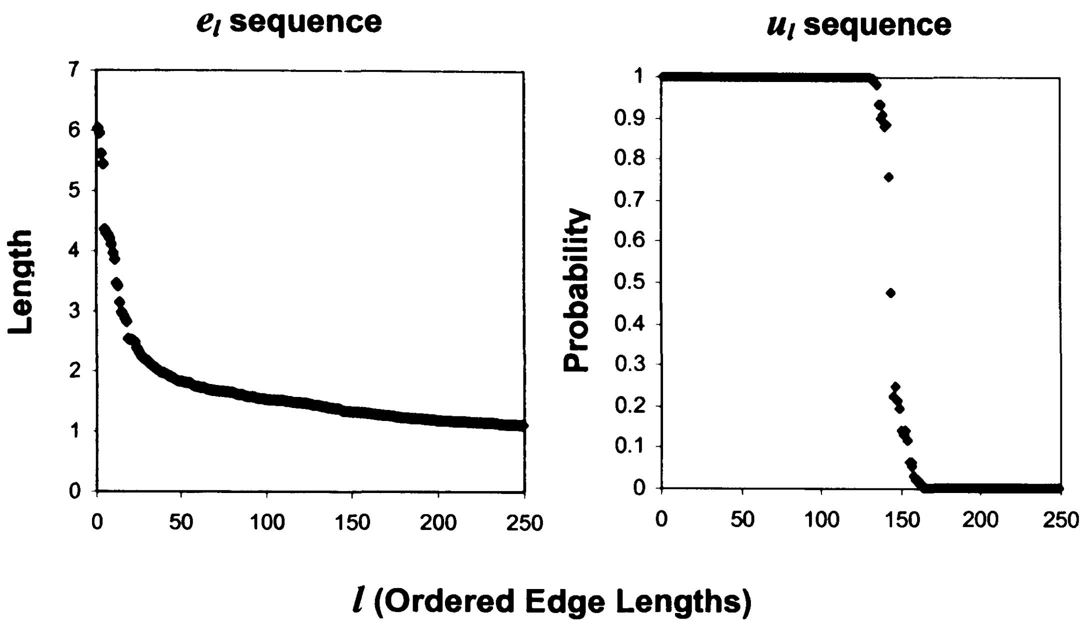
- *Geographic Location*: As, for example, in Gibbons and Machin (2001) who divide the UK into three broad geographical regions or as in Straszheim (1973) who divides San Francisco into 73 geographic zones consisting of one or more census tracts.
- *Property Characteristics*: As, for example, in Allen et al. (1995) who divide properties in Clemson, South Carolina into single-family house, apartment and condominium market segments or Wolverton et al. (1999) who segment the apartment market into one-bedroom, one-bath units; two-bedroom, one-bath units; and two-bedroom, two-bath units.
- *Socioeconomic Characteristics*: As for example in Schnare and Struyk (1976) who stratify the market according to household income levels or Palm (1978) who defines market segments by both income and ethnicity.

We choose clustering variables to reflect all three of these criteria. In particular, property characteristics are described by the property floor area and garden area. Socioeconomic characteristics are defined the four factors indicating the relative wealth, ethnicity, adult age composition and family composition of the property's neighbourhood. Geographic location is accounted for through the use of the spatial smoothing rule described previously.

7.ii. MST initial partition

The data contains 10,870 observations. The MST of the clustering variables for this data was constructed using Mahalanobis distances. Figure 5 presents graphs the first 250 elements of the e-sequence and u-sequence calculated from the edge lengths of the MST.

Figure 5: The e_l-sequence and u_l-sequence of the Minimal Spanning Tree



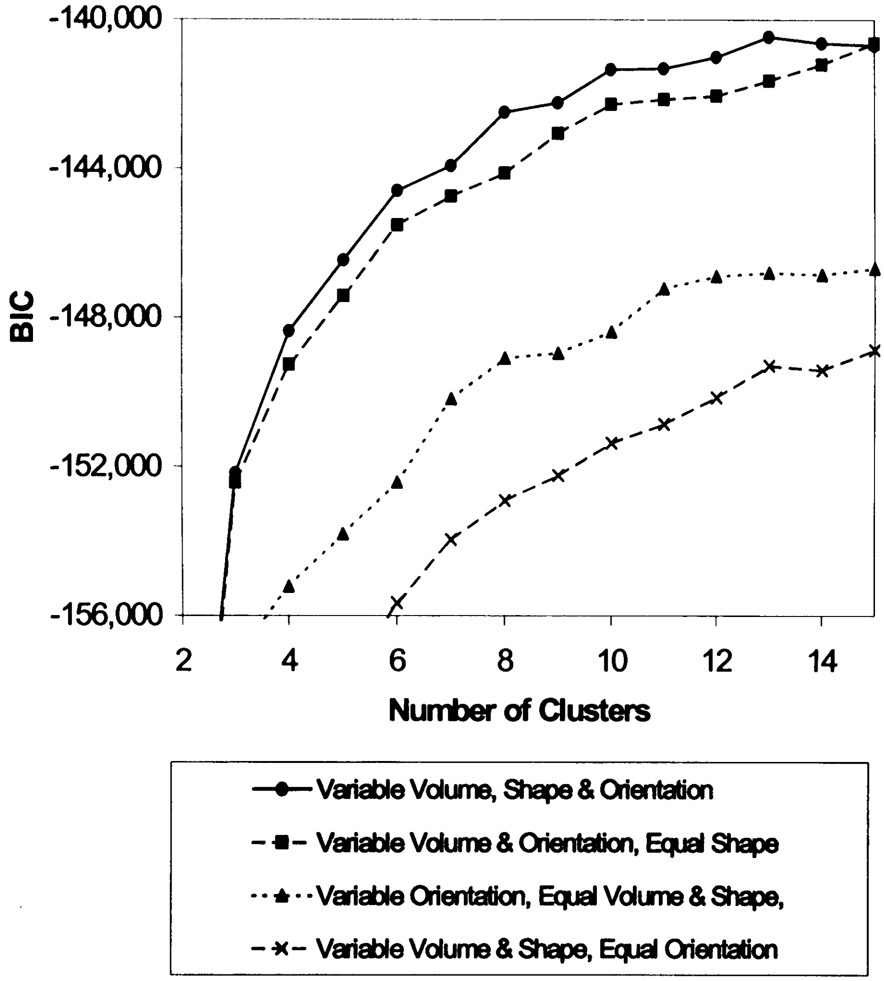
From Figure 5 it is clear that the rate of decline of the e_l sequence reduces significantly after the first 50 longest edge lengths and the u_l sequence stabilises around zero shortly after the 150th longest edge length. Following Posse's (2001) proposition, therefore, we choose to peel the first 175 longest edges of the MST.

Subsequently we prune the MST into 3,830 roughly equal sized clusters. The average cluster size following pruning is 2.83 observations per cluster, the maximum number of observations in any one cluster is 13.

7.iii. Model based clustering of properties with geographical smoothing

Clusters derived from the MST are used to initialise the model-based clustering algorithms. A variety of models corresponding to different numbers of clusters and different cross-cluster restrictions on the cluster covariance matrices were estimated. BIC values for a selection of these models are presented in Figure 6. The four covariance models described in the figure performed significantly better than other possible parameterisations. Indeed, no other model estimated returned BIC scores that would register on this graph.[29]

Figure 6: BIC scores for clustering models assuming different numbers of clusters and different parameterisations of the covariance matrices



From Figure 6 it is clear that the unconstrained model, in which different clusters may differ in size, shape and orientation, outperforms the other models. Again following Posse (2001) we choose the 8 cluster model since this gives a particularly high value for the BIC in the region where the rate of change of the BIC drops significantly.

---

[29] This observation indicates that traditional approaches to clustering such as Ward's (1963) method may be inappropriate since this method is a equivalent to restricting the covariance matrices to be spherical but with differing volumes (Fraley and Raftery, 1998).

Finally, the spatial smoothing algorithm was applied to the clustering solution. After three iterations of the algorithm the classification stabilised, with some 338 properties having changed classification.

7.iv. Market segment descriptions

The statistical procedures described here suggest that the property market in Birmingham is characterised by 8 market segments. Tables 11 and 12 present summary statistics that can be used to compare the characteristics of the properties in the eight market segments. Also Table 13 indicates the type of properties in each market segment as defined by their beacon group classification. We use this data, ward level maps of the concentration of ethnic minority groups (Figure 7) and unemployment levels (Figure 8) and maps of the location of properties (Figures 9 to 16) to interpret and describe the eight market segments.

Figure 7: Ethnic Minority Groups

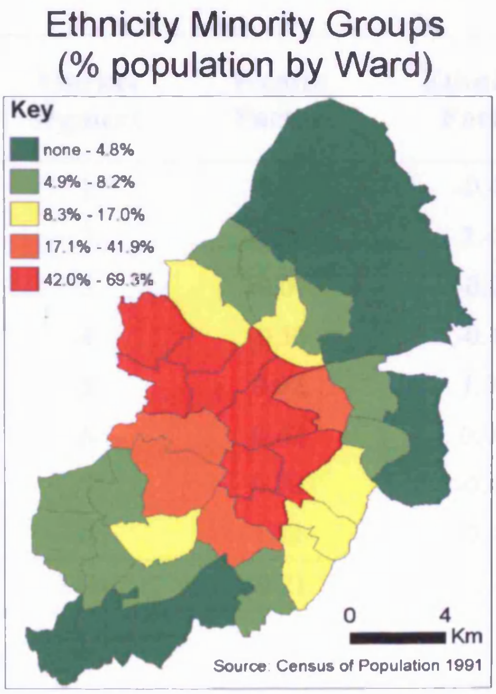


Figure 8: Unemployment

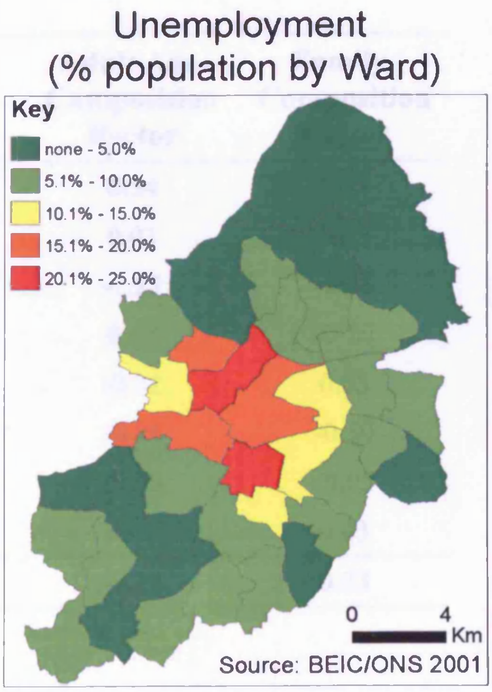


Table 11: Average attribute values for market segments in the Birmingham property market (structural characteristics)

Market segment	Price (£)	Floor Area (m^2)	Garden Area (m^2)	Property Age (yrs before 1997)	% Detached Houses	% Terraced Houses
1	40,564	88	184	63	0.01	0.38
2	31,624	100	85	93	0.00	0.88
3	47,137	97	99	79	0.02	0.73
4	56,307	95	232	52	0.09	0.22
5	53,330	114	221	68	0.12	0.33
6	140,498	192	783	72	0.53	0.08
7	55,806	88	195	58	0.06	0.18
8	99,333	122	397	52	0.40	0.06
Total	59,160	103	227	65	0.12	0.35

Table 12: Average attribute values for market segments in the Birmingham property market (socioeconomic characteristics)

Market segment	Wealth Factor	Ethnicity Factor	Adult Age-Composition Factor	Family-Composition Factor
1	-0.57	-0.47	0.34	0.37
2	-0.36	2.43	0.03	0.83
3	-0.01	-0.06	-1.22	-0.73
4	-0.13	-0.49	0.47	-0.20
5	0.02	1.03	-0.31	0.03
6	0.48	0.08	-0.08	-0.50
7	0.70	-0.42	-0.20	-0.05
8	1.22	-0.49	0.03	0.43
Total	-0.21	0.04	-0.14	0.03

Table 13: Distribution of property types by Beacon Group and market segment

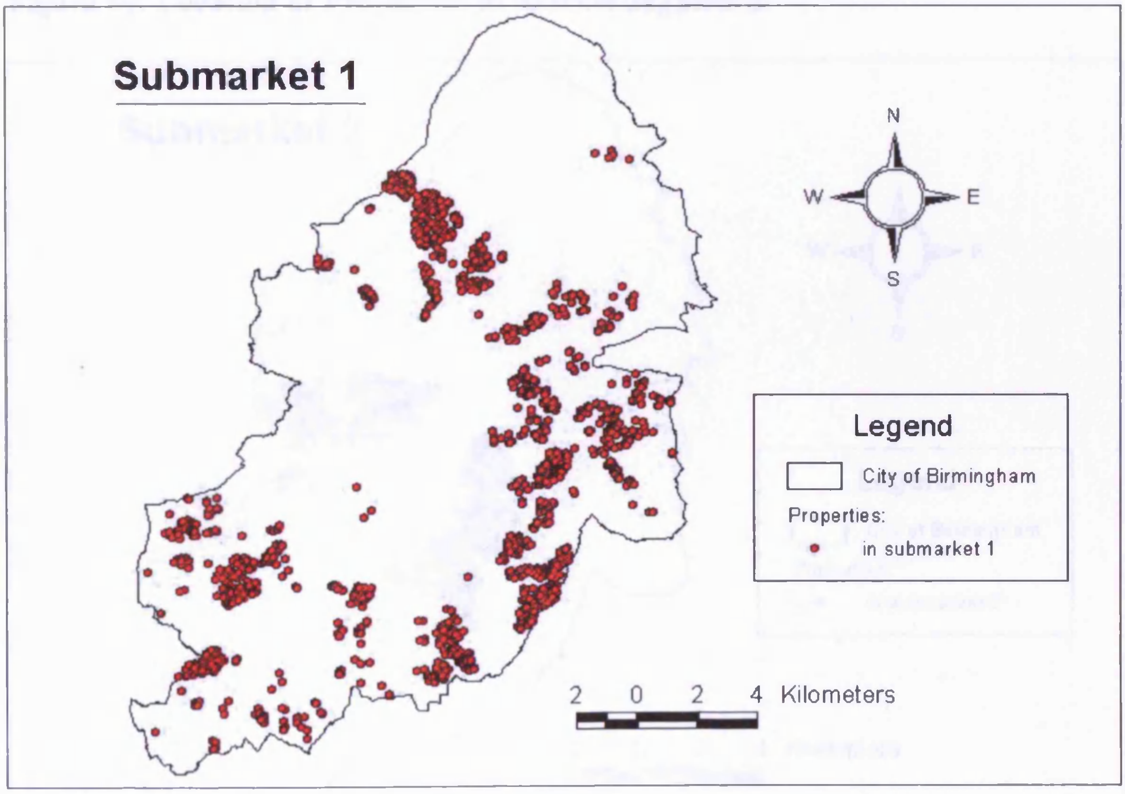
Beacon Group Designation	Market segment							
	1	2	3	4	5	6	7	8
Houses & Bungalows Pre-1919								
Bg1: Unrenovated Cottage			0%	0%			0%	0%
Bg2: Renovated Cottage				0%		1%		0%
Bg3: Small basic "industrial" type	1%	22%	8%	2%	1%		1%	0%
Bg4: Larger "industrial" type	4%	70%	61%	6%	26%	12%	8%	9%
Bg5: Three-storied terrace			0%		4%	4%		0%
Bg8: Small "villa" type	0%	1%	7%	0%	4%	1%	0%	2%
Bg9: Larger "villa" type		0%			1%	14%		2%
Bg10: Large detached houses						8%		0%
Houses & Bungalows 1908-1930								
Bg19: Good quality predate Bg21	0%		0%	1%	2%	3%	0%	4%
Houses & Bungalows 1919-1945								
Bg20: Local Authority "subsidy" type	63%	1%	2%	11%	14%	2%	7%	3%
Bg21: Standard mainly 1930-45	12%	2%	3%	28%	23%	14%	59%	30%
Bg24: Larger dwellings			0%		1%	17%	0%	5%
Bg25: Individually designed						1%		
Houses & Bungalows Post 1945								
Bg30: Local Authority 1945-53	5%	1%	0%	18%	3%	1%	4%	3%
Bg31: Standard post-1953	15%	2%	16%	31%	18%	7%	20%	28%
Bg32: Large dwellings	0%		2%	2%	3%	12%	1%	13%
Bg35: Individually designed						2%		0%
Bg36: "Town Houses" from 1950	0%	0%	0%	1%	1%			0%
Total Dwellings	1,520	1,034	1,549	1,393	1,074	430	2,388	1,460

Market Segment 1: Low Income, White, Ex-Local Authority Housing

Market segment 1 is defined chiefly by the socioeconomic characteristics of households living in the neighbourhoods of these properties. In particular, these neighbourhoods are characterised by a mostly white population, making up some of the poorest communities in Birmingham.

The properties belonging to market segment 1 are relatively small and relatively inexpensive. It is noticeable from the distributional map that there exists a number of large and distinct clusters of properties classified as belonging to this market segment. These clusters correspond to areas of local authority housing. Indeed, almost 70% of the dwellings in this market segment are defined as being ex-local authority owned properties. It seems reasonable to categorise this property market segment as representing a poor, mostly white market segment comprised mainly of ex-local authority housing properties.

Figure 9: Location of Properties in Market Segment 1

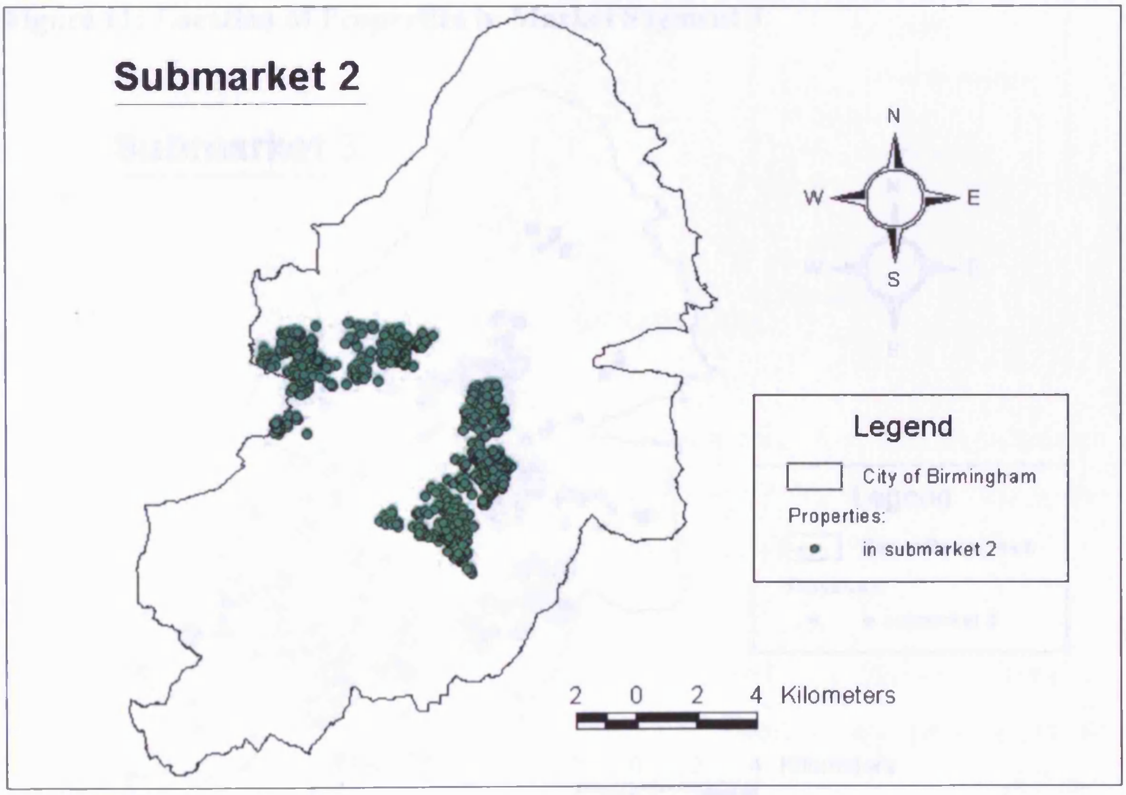


Market Segment 2: Low Income, Black/Asian, Inner-City

Market segment 2 consists of a set of properties concentrated in the inner-city, forming a distinct ring surrounding Birmingham City centre. Once again, the defining feature of this market segment is the socioeconomic characteristics of the neighbourhoods within which the properties are located. These neighbourhoods tend to have high concentrations of residents from the ethnic minorities. Perhaps unsurprisingly, the socioeconomic composition of this market segment is also characterised by relative poverty, a wide range of adult ages and a high level of households with children.

The properties in this market segment are, for the most part, turn of the century terraced houses; 92% are classified in these Beacon Groups. On average these properties command the lowest prices of any market segment. We characterise this market segment as being relatively small properties in the inner-city, located in poor Black or Asian neighbourhoods.

Figure 10: Location of Properties in Market Segment 2



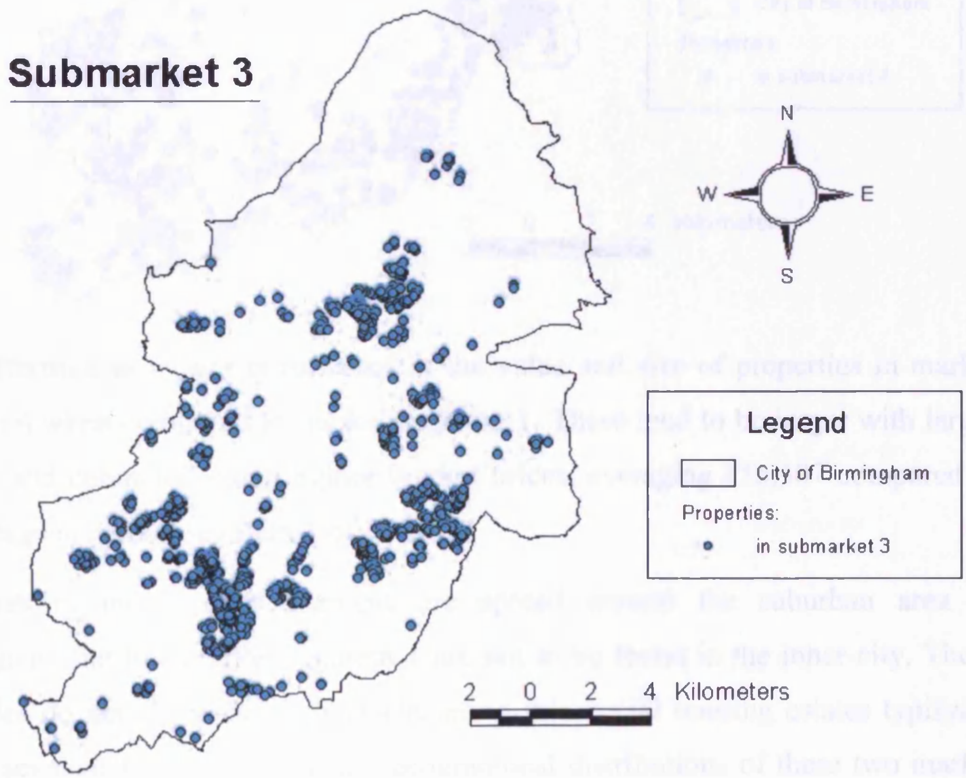
Market Segment 3: Young Adults, Without Children (First-Time Buyers?)

Physically, properties classified as belonging in market segment 3 show many similarities to those falling in market segment 2; they tend to be small turn of the century terrace properties with relatively small gardens. In contrast, on average properties in this market segment command a selling price some £16,000 greater than those in market segment 2.

Similarly, this market segment is relatively wealthier than market segment 2 and has a much more mixed ethnicity. The defining feature of this market segment is that it comprises neighbourhoods inhabited by young adults without children.

We surmise from this fact that market segment 3 is a young, possibly first-time buyers market. The geographic distribution of properties in this market segment accords with this interpretation. Properties in market segment 3 form clusters falling in a wide band surrounding the inner city, within easy commuting distance of the city centre. Further, a particularly large concentration of properties to the south and west of the city centre is located around the University and Hospital complex.

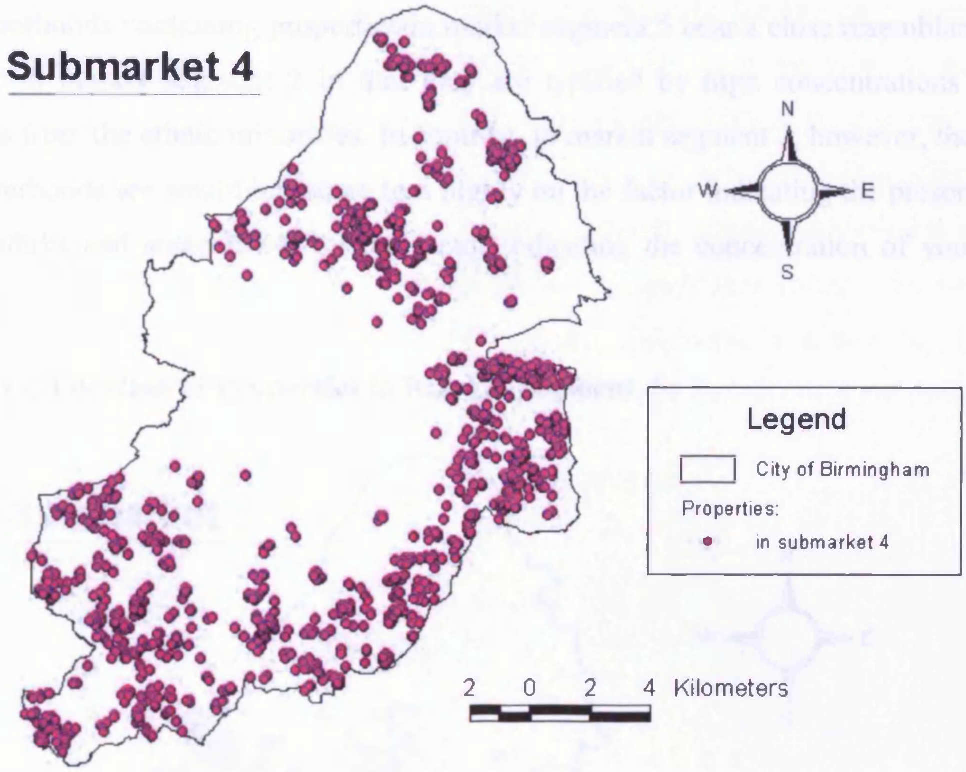
Figure 11: Location of Properties in Market Segment 3



Market Segment 4: Middle Income, White

Neighbourhoods containing properties characterised as belonging to market segment 4 show many similarities to those in market segment 1. They are dominated by an ethnically white population and tend to have a relatively old population. In contrast to market segment 1, however, residents tend to be relatively more wealthy (though not as wealthy as those in market segments 6 to 8) and have fewer families with children.

Figure 12: Location of Properties in Market Segment 4



This difference in wealth is reflected in the value and size of properties in market segment 4 when compared to market segment 1. These tend to be larger with larger gardens and command much higher market prices, averaging £56,307 compared to the average in market segment 1 of £40,564.

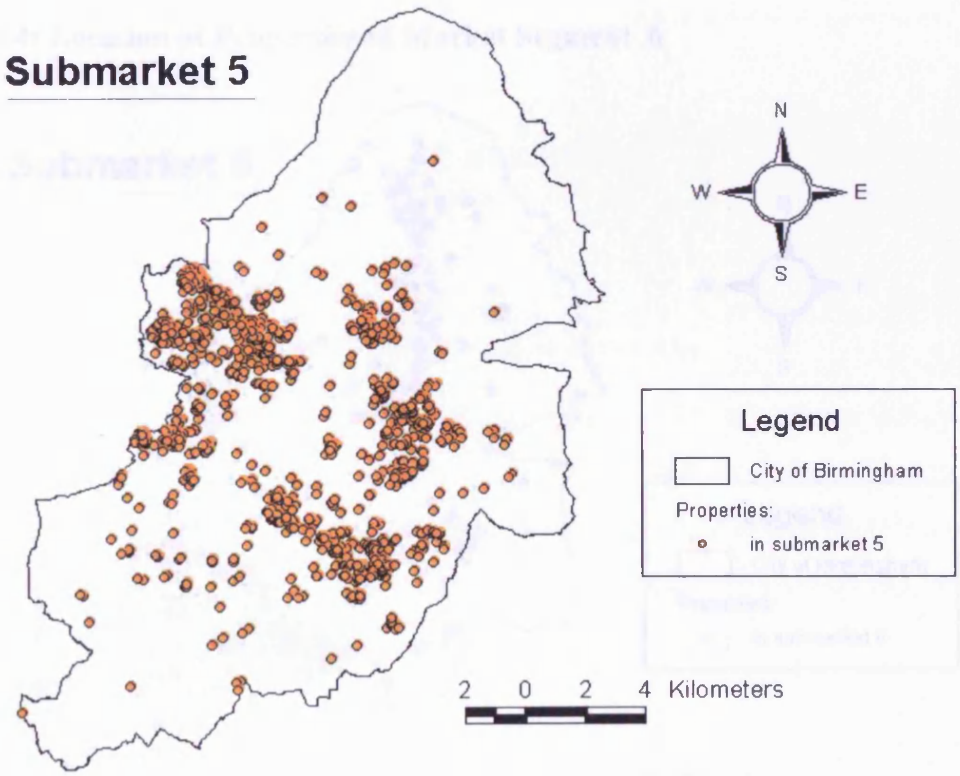
Properties in this market segment are spread around the suburban area of Birmingham but like market segment 1 are not to be found in the inner-city. These properties do not show the distinct clustering into social housing estates typifying market segment 1. Indeed when the geographical distributions of these two market segments are compared it can be seen that frequently properties classified as belonging to market segment 4 fall in loose circles at the periphery of social housing estates. Further, almost 60% of properties in this market segment are “standard” dwellings according to their beacon group classification.

We interpret this market segment as being a middle income, ethnically white market segment. Properties in this market segment tend to be of a standard type and are dispersed across the suburban area with particularly strong concentrations in areas bordering local authority housing estates.

Market Segment 5: Middle Income, Black/Asian

Neighbourhoods containing properties in market segment 5 bear a close resemblance to those in market segment 2 in that they are typified by high concentrations of residents from the ethnic minorities. In contrast, to market segment 2, however, these neighbourhoods are wealthier, score less highly on the factor indicating the presence of old adults and score lower on the factor indicating the concentration of young families.

Figure 13: Location of Properties in Market Segment 5

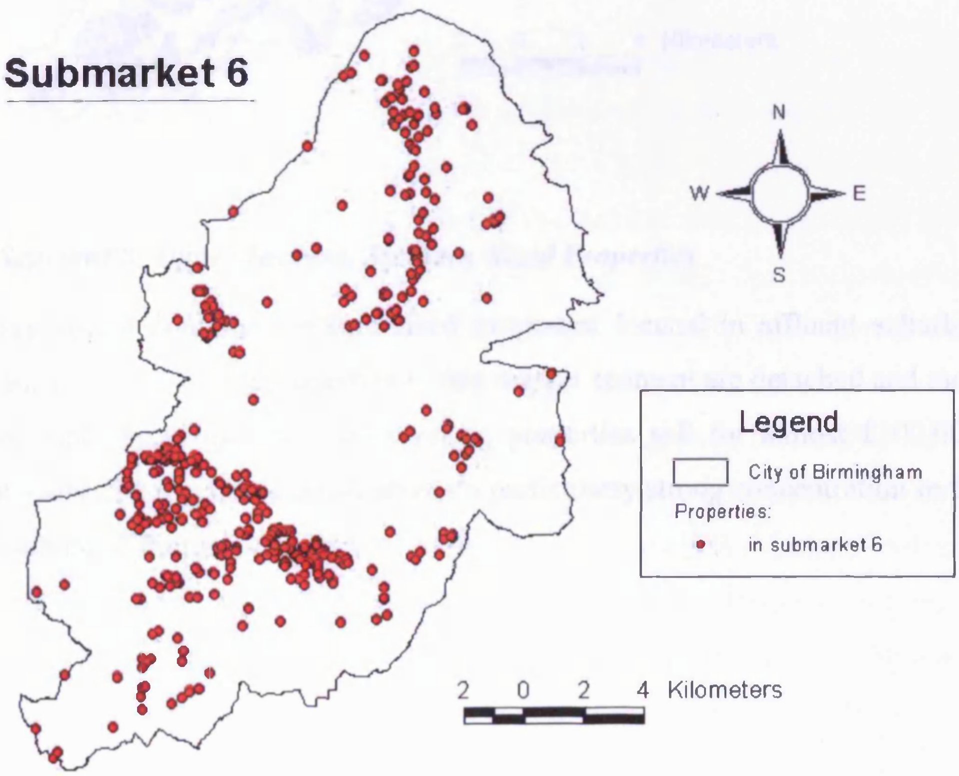


Properties in market segment 5 are relatively larger than those in market segment 2 with larger gardens and commanding market prices some £25,000 greater. Like market segment 4, almost 70% of properties are of a "standard" type. Geographically, properties fall in a broad ring surrounding the relatively poorer Black and Asian communities inhabiting the core of the inner city. Notice, that there is little dispersion of market segment 5 into suburban areas. It appears that this market segment represents middle income Black and Asian communities.

Market Segment 6: Upper Income, Large Properties

The remaining three market segments represent properties located in wealthier suburban locations. Indeed, properties in market segment 6 are amongst the most expensive in Birmingham averaging £140,498. This price tag reflects the fact that the properties in this market segment are very large, mostly detached suburban properties. Geographically these are quite widely scattered, though properties in this market segment are have particular concentrations in a band to the south and west of the city, in the wealthy neighbourhoods of Harborne, Edgbaston and Moseley, and also in the northern suburbs of the city around the fringes of Sutton Coldfield Park.

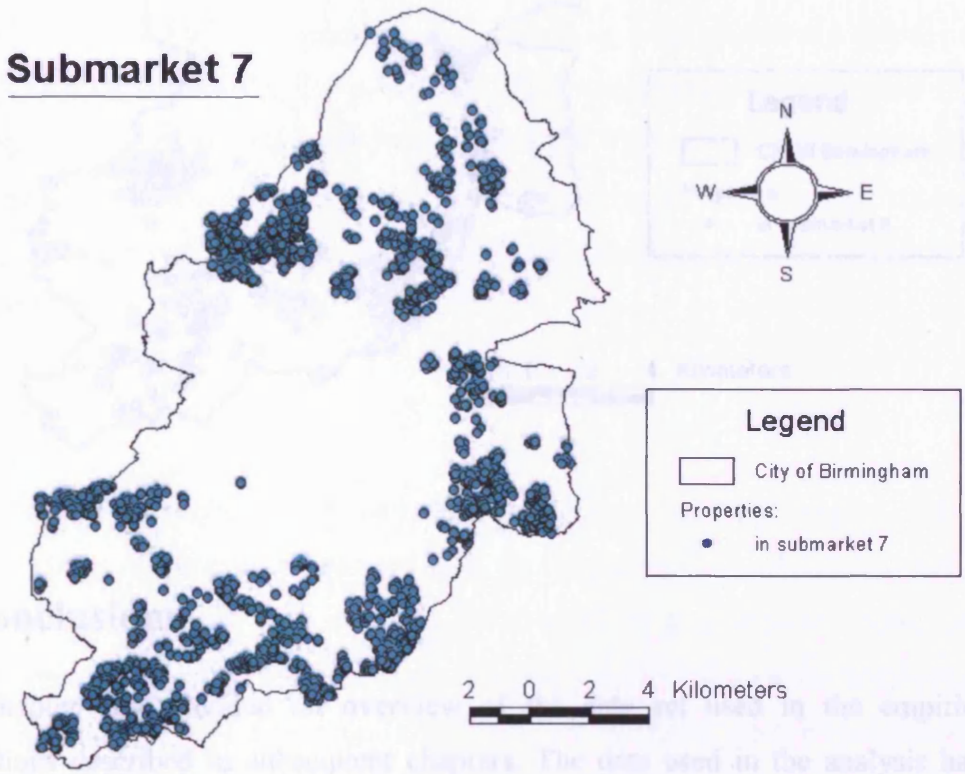
Figure 14: Location of Properties in Market Segment 6



Market Segment 7: Upper Income, Small Properties

Many similarities exist between properties in market segment 7 and those classified as belonging to market segment 4. They are generally small with medium sized gardens and command a price tag of around £56,000. In contrast, properties in market segment 4 are located in affluent neighbourhoods.

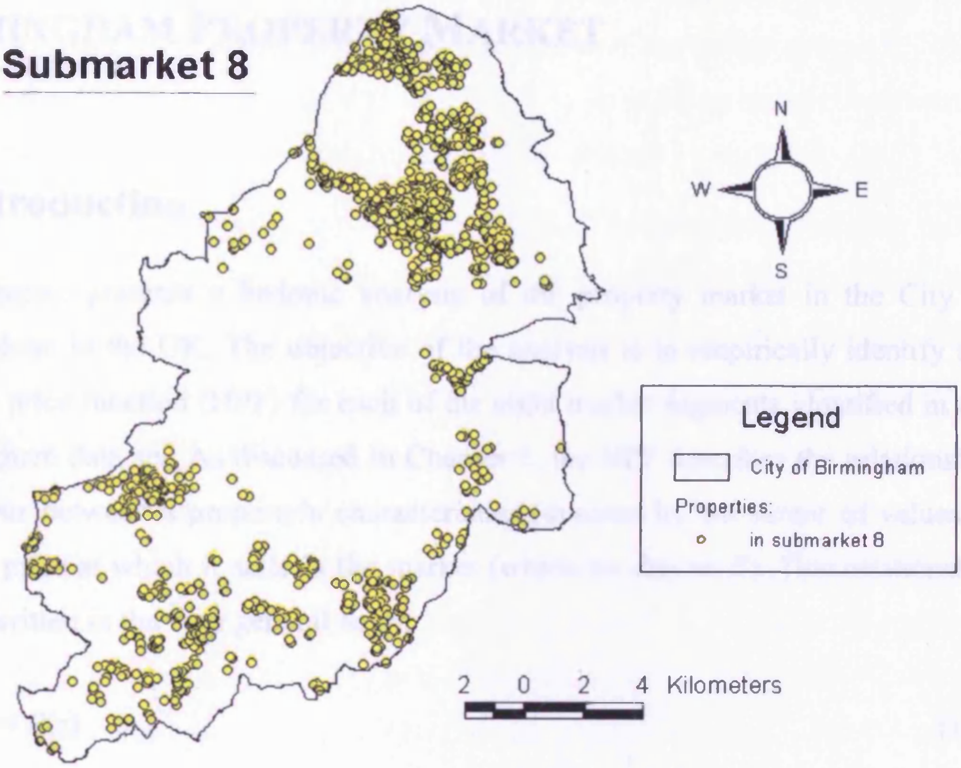
Figure 15: Location of Properties in Market Segment 7



Market Segment 8: Upper Income, Medium-Sized Properties

Market segment 8 contains medium-sized properties located in affluent suburban areas of Birmingham. Many properties in this market segment are detached and most have reasonably large gardens. On average, properties sell for almost £100,000. Geographically, the market segment shows a particularly strong concentration in the northern suburb of Sutton Coldfield.

Figure 16: Location of Properties in Market Segment 8



8. Conclusions

This Chapter has provided an overview of the data set used in the empirical applications described in subsequent chapters. The data used in the analysis have been collected from a large number of sources and collated with aid of GIS. Two major pre-processing tasks have been carried out on the data. First, using factor analytical techniques the numerous variables describing the socioeconomic characteristics of neighbourhoods have been condensed into four factors each capturing a major dimension of similarity and difference in neighbourhood composition. Second, model-based clustering has been applied to identify market segments in the Birmingham property market. The model-based clustering algorithms were initialised from an efficient initial partition of the data derived from the MST of the data. The cluster analysis suggests that a good candidate solution is provided by an 8 cluster partition. Inspecting the property and neighbourhood characteristics of each cluster reveals that this partition gives rise to distinct and easily interpretable market segments.

CHAPTER 5: A HEDONIC ANALYSIS OF THE CITY OF BIRMINGHAM PROPERTY MARKET

1. Introduction

This Chapter presents a hedonic analysis of the property market in the City of Birmingham in the UK. The objective of the analysis is to empirically identify the hedonic price function (HPF) for each of the eight market segments identified in the Birmingham data set. As discussed in Chapter 1, the HPF describes the relationship that exists between a property's characteristics (denoted by the vector of values z) and the price at which it sells in the market (which we denote P). This relationship can be written in the very general form;

$$P = P(z) \tag{1}$$

The research reported in this Chapter represents one of the most detailed hedonic pricing studies carried out to date. The data set that forms the basis of this study is perhaps the richest of its kind yet to be constructed for any property market. The model-based clustering techniques used to identify market segments represent the state-of-the-art in data partitioning. In addition, the empirical analysis presented in this Chapter breaks new ground in applied econometric research.

In specifying an empirical model of the HPF, researchers are faced by a number of important questions. In Section 2 we deal with the thorny issue of selecting functional form. On the whole, the less explicit the researcher is in specifying the relationship between the characteristics data and property prices, the less likely it is that the model will be mis-specified. In our application we employ a semiparametric estimation method known as the *Partially Linear Model* (PLM) (Robinson, 1988) which imposes relatively few assumptions on the relationship between prices and property characteristics.

Despite the comprehensive nature of the Birmingham data set, it is inevitable that some location specific influences on property prices are not accounted for by the variables in the regression equation. In Section 3 we discuss two approaches for dealing with missing spatial covariates. The first approach is to simply include numerous region specific *locational constants*. The second approach generalises the locational constants through the application of a *spatial smoothing estimator* (SSE). In effect, the SSE uses nonparametric techniques to estimate a separate observation-specific locational constant for each property in the dataset.

Previous hedonic analyses have employed the PLM to introduce flexibility in the relationship between prices and characteristics (e.g. Anglin and Gencay, 1996) and the SSE to account for omitted locational covariates (Gibbons and Machin, 2002; Gibbons, 2002), however, this is the first application in which both have been combined in a single model.

Whilst the use of locational constants or the SSE might effectively capture important features of location not accounted for by the covariate data, there is every likelihood that many minor features of location will result in spatial correlation in the nuisance process. Section 4 describes a a general method of moments (GMM) estimator proposed by Kelejian and Prucha (1999) that accounts for spatial autocorrelation so as to provide robust estimates of the parameters of the HPF.. Again this is the first time that semiparametric methods that introduce flexibility into the specification of the HPF have been coupled with the Kelejian and Prucha GMM estimator to provide robust estimates of the parameters of a HPF.

The estimation strategy followed in the analysis is summarised in Section 5. Finally, Section 6 provides details of the regression results, reporting estimates of the parameters of the HPF for each of the property market segments and deriving estimates of the implicit prices for traffic-related noise pollution.

2. Functional Form

The objective of the empirical investigation of the Birmingham data set is to provide an empirical estimate of the HPF (Equation 1). To do so, we define a regression equation with an additive error term that, in general form, can be written;

$$g(P_i) = h(z_i) + \varepsilon_i \tag{2}$$

where P_i is the price of observation i, z_i is a row vector of associated property characteristics and ε_i is an observation specific error term. The two functions $g(\cdot)$ and $h(\cdot)$ determine the exact nature of the relationship between the dependent variable (P_i) and explanatory variables (z_i). To avoid complicating the notation we ignore the fact that a separate regression equation is estimated for each market segment.

Our first task in estimating the parameters of the HPF is to specify functional forms for $g(\cdot)$ and $h(\cdot)$. Unfortunately, economic theory provides little guidance on the nature of the relationship between property prices and property characteristics. In general, researchers have opted to define $g(\cdot)$ as the log transformation. We follow that convention here as this transformation confers at least two advantages.

- First, the distribution of property prices in a market tends to show considerable right skew. Such data distributions are often associated with heteroskedasticity and/or non-normality of errors, both of which complicate estimation. As illustrated for the Birmingham data in Figure 1, transforming to a log scale may reduce such problems by generating a more symmetrical distribution of the dependent variable.
- Second, using a log transformation allows for readily interpretable coefficient estimates. For example, the coefficient on a regressor entered in simple linear form indicates the constant percentage response in property price to a one unit absolute increase in the regressor.

The regression model (2) can be rewritten as;

$$\ln P_i = h(z_i) + \varepsilon_i \tag{3}$$

Estimating a model such as (3) often requires the imposition of strong assumptions on the function $h(\cdot)$. For example, many studies assume that this is a linear function, giving rise to the familiar semi-log model;

$$\ln P_i = z_i \gamma + \varepsilon_i \tag{4}$$

where γ is a column vector of parameters.

Figure 1: Distribution of price data plotted on a log scale

Of course, little in economic theory would suggest that such a strong assumption is valid. As a result, considerable attention has been focused on the use of more flexible specifications. In particular, a number of researchers have investigated the use of the Box-Cox flexible functional form (e.g. Cropper, Deck and McConnell, 1988; Cheshire and Sheppard, 1998). Whilst, this approach allows the regression model to more accurately reflect the patterns of association inherent in the data, it also has a number of drawbacks (as discussed by Ramussen and Zuehlke; 1990).

An alternative to increasing the degree of parameterisation of the regression model is to adopt a non-parametric regression approach. Here, the function $h(\cdot)$ is dictated entirely by the data ensuring the regression function is extremely robust to misspecification. Unfortunately, non-parametric estimation of $h(\cdot)$ is only realistic when z_i contains only a few regressors. When there are many regressors, non-parametric response coefficients may be very imprecise.

An intermediate strategy is to employ a semiparametric form such as that proposed by Robinson (1988). Here part of the model is specified parametrically whilst the rest is estimated using non-parametric techniques. Robinson's model is of the form

$$\ln P_i = z_i \beta + q(x_i) + \varepsilon_i \tag{5}$$

where z_i is a k-vector of regressors associated with a k-vector of parameters $\boldsymbol{\beta}$, whilst $\boldsymbol{x}_i$ is a p-vector of regressors whose influence on property prices is determined by the unknown function $q(\cdot)$. We call this a *Partially Linear* (PL) model.

Robinson shows that the model in Equation (5) can be rewritten as;

$$\ln P_i - E[\ln P \mid x_i] = (z_i - E[z \mid x_i])\boldsymbol{\beta} + \varepsilon_i \tag{6}$$

suggesting that $\boldsymbol{\beta}$ can be estimated in a two-step procedure;

- First, the unknown conditional means $E[\ln P \mid x_i]$ and $E[z \mid x_i]$ are estimated using a non-parametric estimation technique.
- Second, the estimates are substituted in place of the unknown functions in Equation (6) and ordinary regression techniques employed to estimate $\boldsymbol{\beta}$.

Indeed, Robinson shows that the resulting parameter estimates are asymptotically equivalent to those that would be derived if the true functional form of $q(\cdot)$ were known and could be used in the estimation. Robinson's model was pioneered in the hedonic literature by Anglin and Gencay (1996).

In this case, the semiparametric specification has a number of advantages;

- First, the structural characteristics of a property are likely to be most important in determining a property's market price. The empirical model would be considerably more robust if these variables were included in the unknown function $q(\cdot)$. Referring to the data summaries provided in Table 10 of Chapter 4, it is clear that the majority of these characteristics are defined as dummy variables. As Anglin and Gencay (1996) discuss the inclusion of dummy variables in $q(\cdot)$ effects the scale but not the curvature of that function. A reasonable approximation, therefore, would be to include these dummy variables in the linear part of the model. Of the remaining structural variables, those defining a property's floor area, garden area and age, provide a reasonably accurate picture of a property's structure (age proxying for both quality characteristics and architectural design features).

- Second, the variables of interest in this research project are those describing a property's exposure to noise. Including these in the linear part of the model allows for ease of interpretation of parameter estimates, simplifies the calculation of implicit prices and facilitates comparison of estimates with other studies.

Accordingly the vector $\boldsymbol{x}$ entering the unknown function $q(\cdot)$ consists of the log of a property's floor area, the log of a property's garden area and the property's age. Further, property prices in the UK have been reasonably volatile over recent decades. To account for price movements over the course of 1997 a continuous variable indicating the date of sale is included in the function $q(\cdot)$. The choice of variables to be included in the nonparametric function $q(\cdot)$ is summarised in Table 1.

Table 1: Variables included in nonparametric part of the hedonic price regression

Variable	Description
Area	Natural logarithm of property floor area in m^2
Garden	Natural logarithm of garden size in m^2
Age	Age of property in decades previous to 1997
Sale Date	Date of sale in days from 1st January 1997

Following Anglin and Gencay (1998) $E[\ln P \mid \boldsymbol{x}_i]$ and $E[z \mid \boldsymbol{x}_i]$ are estimated using non-parametric kernel regression. The kernel estimator of the density of the random vector $\boldsymbol{x}$ evaluated at $\boldsymbol{x}_i$ is given by;

$$\hat{f}_H(\boldsymbol{x}_i) = \frac{1}{N}\sum_{j=1}^{N} K_H(\boldsymbol{x}_j - \boldsymbol{x}_i) \tag{7}$$

where $K_H(\boldsymbol{x}_j - \boldsymbol{x}_i) = \det(\boldsymbol{H})^{-1} K(\boldsymbol{H}^{-1}(\boldsymbol{x}_j - \boldsymbol{x}_i))$ for some multivariate kernel function $K(\boldsymbol{u})$ and for a given $p \times p$ matrix of bandwidths, $\boldsymbol{H}$.

In effect, the kernel density estimator counts the number of observations in the dataset in close proximity to $\boldsymbol{x}_i$. The density at $\boldsymbol{x}_i$ is approximated by dividing this count by the number of observations in the dataset. Whether observations $\boldsymbol{x}_j$ are considered close to $\boldsymbol{x}_i$ is determined by the bandwidth matrix $\boldsymbol{H}$. The larger the elements of the bandwidth matrix, the more observations are drawn into the count. Further the weight allotted to each observation in the count is determined by the kernel function $K(\boldsymbol{u})$. The kernel function must be symmetric, continuously differentiable and integrate to unity. Moreover, most commonly used kernel functions allot greater weight to observations in close proximity to $\boldsymbol{x}_i$ than those further away.

Here we use a matrix of bandwidths determined by $\boldsymbol{S}$, the sample covariance matrix of $\boldsymbol{x}$, such that;

$$\boldsymbol{H} = h\boldsymbol{S}^{\frac{1}{2}} \tag{8}$$

for some positive scalar h. In this case, the argument to the kernel function can be written;

$$\boldsymbol{u} = h^{-1}\boldsymbol{S}^{-\frac{1}{2}}\left(\boldsymbol{x}_j - \boldsymbol{x}_i\right) \tag{9}$$

Further, we employ a multivariate Gaussian Kernel of the form;

$$K(\boldsymbol{u}) = (2\pi)^{-\frac{P}{2}}\exp\left(-\frac{1}{2}\boldsymbol{u}'\boldsymbol{u}\right) \tag{10}$$

As such the kernel density estimator of Equation (7) can be written in the specific form;

$$\hat{f}_h(\boldsymbol{x}_i) = \frac{1}{N}\sum_{j=1}^{N} h^{-P}\det(\boldsymbol{S})^{-\frac{1}{2}}(2\pi)^{-\frac{P}{2}}\exp\left(-\frac{1}{2}\boldsymbol{u}'\boldsymbol{u}\right) \tag{11}$$

To generalise notation, let r represent the element whose conditional expectation we wish to estimate. In our case, therefore, r denotes any element of the z vector or the log of property price. Then the conditional expectations we wish to estimate are given by;

$$E[r \mid x_i] = \frac{\int r\, f(r, x_i)\, dy}{\int f(r, x_i)\, dy} \tag{12}$$

Nadaraya-Watson kernel regression estimates (12) by replacing the numerator and denominator with their equivalent kernel density according to;

$$\hat{E}_h[r \mid x_i] = \frac{\sum_{j=1}^{N} r\, K_h(x_j - x_i)}{\hat{f}_h(x_i)} \tag{13}$$

A central issue in nonparametric estimation is the choice of bandwidth, ***h***. The bandwidth parameter determines the degree of smoothing of the function $\hat{E}_h[r \mid x_i]$. Too large a value for ***h*** induces bias and too small a value results in imprecise estimates.

Once again, following Anglin and Gencay (1998) we select bandwidths using a data driven technique known as *cross-validation*. As they point out, a seemingly natural way to select ***h*** is to choose the bandwidth that minimises the sum of squared residuals from the regression equation;

$$MSE = n^{-1} \sum_{i=1}^{N} (\ln P_i - z_i \boldsymbol{\beta} - q_h(x_i))^2 \tag{14}$$

where the estimator of $q_h(x_i)$ is given by;

$$\hat{q}_h(x_i) = \frac{\sum_{j=1}^{N} (\ln P_i - z_i \hat{\boldsymbol{\beta}}) K_h(x_j - x_i)}{\hat{f}_h(x_i)} \tag{15}$$

Unfortunately, this procedure falls down because the objective function, *MSE*, reduces to zero for any ***h*** smaller than the closest two data points in the sample. For such values of h the conditional mean function given by $q_h(\boldsymbol{x}_i)$ puts all weight on the ith observation such that $q_h(\boldsymbol{x}_i)$ perfectly predicts $\ln P_i$.

Accordingly, the criterion function in (15) cannot be used to decide upon the optimal bandwidth. Rather researchers employ the cross-validation statistic;

$$MSE_{CV} = n^{-1}\sum_{i=1}^{N}\left(\ln P_i - z_i\boldsymbol{\beta} - q_{h,i}(\boldsymbol{x}_i)\right)^2 \tag{16}$$

The cross-validation statistic avoids the problems of the raw MSE statistic by employing a conditional mean function $q_{h,i}(\boldsymbol{x}_i)$ that is calculated by *leaving out* the i^{th} observation;

$$\hat{q}_{h,i}(\boldsymbol{x}_i) = \frac{\sum_{j\neq i}^{N}\left(\ln P_i - z_i\hat{\boldsymbol{\beta}}\right)K_h(\boldsymbol{x}_j - \boldsymbol{x}_i)}{\sum_{j\neq i}^{N}K_h(\boldsymbol{x}_j - \boldsymbol{x}_i)} \tag{17}$$

The cross-validation procedure requires a grid search for optimal h. The regression equation (6) is re-estimated numerous times using different values of h. For each value of h the cross-validation statistic is estimated using (16) and the h providing the minimum value for this statistic is chosen as the optimal bandwidth.

Here we improve on the estimation procedure of Anglin and Gencay (1998) by using an adaptive kernel estimator. The motivation behind the adaptive kernel estimator is to improve estimation of the conditional expectation functions $E[\ln P \mid \boldsymbol{x}_i]$ and $E[z \mid \boldsymbol{x}_i]$ by allowing the bandwidth to vary with the density of $\boldsymbol{x}$. Thus where data is relatively sparse the adaptive kernel uses a relatively wide bandwidth, whilst when data is abundant the bandwidth is commensurately reduced.

Adaptive kernel estimation requires a two-stage estimation procedure. First a pilot bandwidth, h_p, is employed to estimate the density of $\boldsymbol{x}$; $\hat{f}_{h_p}(\boldsymbol{x})$. Using this estimated density we calculate;

$$\lambda_i = \left(\frac{\hat{f}_{h_P}(\boldsymbol{x}_i)}{\eta} \right)^{-\rho} \tag{18}$$

where η is a normalisation factor given by $\ln \eta = \sum_j \ln \hat{f}_{h_P}(\boldsymbol{x}_j) / n$ and ρ is a parameter taking a value between 0 and 1 chosen by the researcher. Here we select a value for ρ of 0.25.

In the second stage, the adaptive kernel generalises (13) by using a new observation specific bandwidth parameter $h_j = h\lambda_j$ to estimate the conditional expectations.

Given the cross-validated, adaptive kernel estimates of $E[\ln P \mid \boldsymbol{x}_i]$ and $E[z \mid \boldsymbol{x}_i]$, we are left with the task of estimating the semiparametric model;

$$\ln P_i - E[\ln P \mid \boldsymbol{x}_i] = (z_i - E[z \mid \boldsymbol{x}_i])\boldsymbol{\beta} + \varepsilon_i \tag{6}$$

To maintain clarity, let us introduce some new notation. Let tilde indicate differences from non-parametric expectations, such that $\widetilde{y}_i = \ln \widetilde{P}_i = \ln P_i - E[\ln P \mid \boldsymbol{x}_i]$ and $\widetilde{z}_i = z_i - E[z \mid \boldsymbol{x}_i]$. Consequently, Equation (6) simplifies to;

$$\widetilde{y}_i = \widetilde{z}_i \boldsymbol{\beta} + \varepsilon_i \tag{19}$$

One possibility is to estimate (19) using (no constant) ordinary least squares (OLS) according to the familiar formula;

$$\hat{\boldsymbol{\beta}}^{SC} = \left(\widetilde{\boldsymbol{Z}}' \widetilde{\boldsymbol{Z}} \right)^{-1} \widetilde{\boldsymbol{Z}}' \widetilde{\boldsymbol{Y}} \tag{20}$$

where $\widetilde{\boldsymbol{Z}}$ is the $N \times k$ matrix of data formed by stacking the $\widetilde{z}_i$ vectors and $\widetilde{\boldsymbol{Y}}$ is the N x 1 vector with elements $\widetilde{y}_i$. For reasons that will shortly become clear we label the parameter estimates defined by (20) as $\hat{\boldsymbol{\beta}}^{SC}$, where SC stands for *Spatial Constants*.

3. Unaccounted Spatial Variation in Property Prices

Up to this point, our statistical analysis has ignored the spatial organisation of the data. In effect, we have assumed that the observations of property sales are independent such that we can glean no information on the selling price of a property from the selling price of other properties. Of course, this is hardly likely to be the case. Properties that are located near to each other in space are also likely to share common environmental, accessibility, neighbourhood and perhaps even structural characteristics. Even once we account for the values of known covariates, omitted variables are likely to induce spatial correlation amongst property prices. Since we know the spatial distribution of properties, it is possible to model this unaccounted spatial variation. Here we contrast two different approaches.

Spatial Constants Model (PLSC)

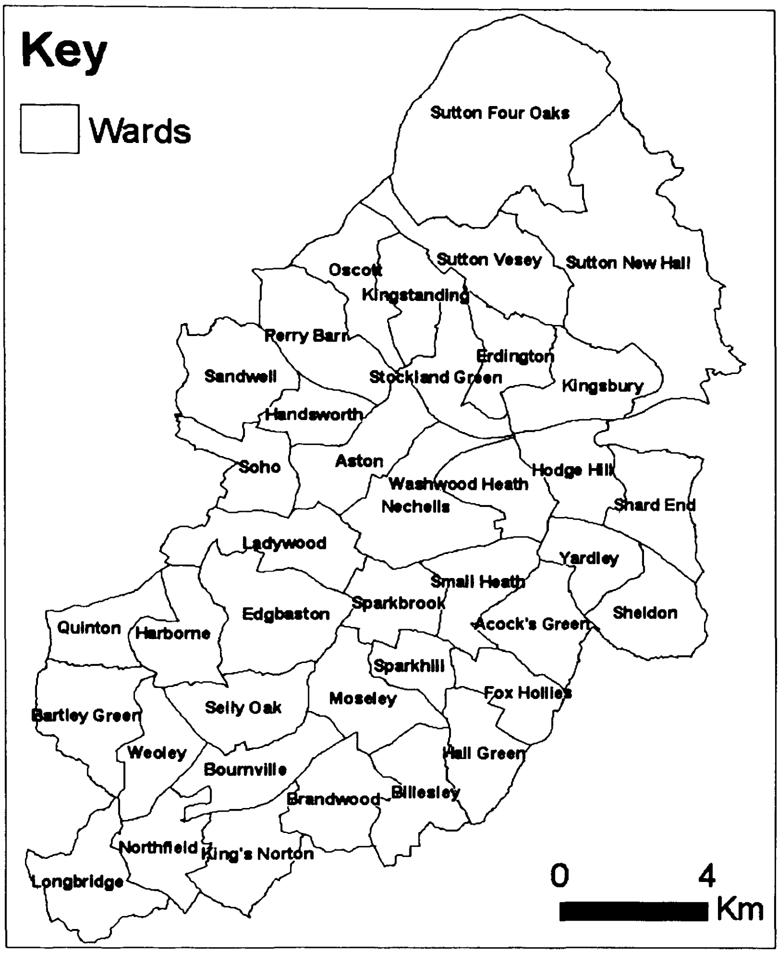
The simpler of these two approaches is to partition the Birmingham cityscape into a patchwork of areas. Ideally these areas should reflect regions of relative physical and socioeconomic homogeneity. Here we choose to use the political boundaries that define the 39 electoral wards in the City of Birmingham (see Figure 2). As such the covariate data is augmented with a set of spatial constants indicating location in one of these 39 wards.

Hedonic functions for each market segment are estimated by selecting one ward as the baseline and including spatial constants for the remaining wards in the covariate vector z_i. The parameters of the *PL model with spatial constants* (PLSC) are estimated by applying Equation (20). Parameters for the spatial constants identify the average difference in property prices between each ward and the baseline ward once the influence of all the other covariates has been accounted for.

Spatial Smoothing Model (PLSS)

Of course, including a set of spatial constants to capture similarities in the prices of closely located properties is relatively crude. For example, this model assumes that a property on the boundary of a ward has more in common with the other properties in its own ward than properties in its direct vicinity but on the other side of the ward boundary.

Figure 2: Wards of the City of Birmingham



An alternative approach has been suggested by Gibbons and Machin (2002) and Gibbons (2002). Their *smooth spatial effects* (SSE) estimator is a simple extension of the PL model described above.

The location of a property can be expressed by the vector $\boldsymbol{c}_i = (c_{1i}, c_{2i})$ where c_{1i} is an easting and c_{2i} is a northing. Gibbons and Machin (2002) define their SSE estimator as a PL model in which the regressors to be modelled nonparametrically, $\boldsymbol{x}_i$, are simply the property locations, $\boldsymbol{c}_i$. Roughly speaking, the smooth spatial effects estimator strips location-specific determinants of property prices out of the regression equation.

Here we extend the Gibbons and Machin model by including both locational and property characterisitcs data in the nonparametric part of the PL model. We call this a *PL model with Spatial Smoothing* (PLSS). The PLSS model reformulates (5) as;

$$\ln P_i = z_i\beta + q(\boldsymbol{x}_i, \boldsymbol{c}_i) + \varepsilon_i \tag{21}$$

where $\boldsymbol{x}_i$ is the vector of variables listed in Table 1. Similarly, (6) is reformulated as;

$$\ln P_i - E[\ln P \mid \boldsymbol{x}_i, \boldsymbol{c}_i] = (z_i - E[z \mid \boldsymbol{x}_i, \boldsymbol{c}_i])\boldsymbol{\beta} + \varepsilon_i \tag{22}$$

The quantities $E[\ln P \mid \boldsymbol{x}_i, \boldsymbol{c}_i]$ and $E[z \mid \boldsymbol{x}_i, \boldsymbol{c}_i]$ are once again estimated using non-parametric kernel regression according to;

$$\hat{E}_{h,b}[r \mid \boldsymbol{x}_i, \boldsymbol{c}_i] = \frac{\sum_{j=1}^{N} r \; K_h(\boldsymbol{x}_j - \boldsymbol{x}_i) K_b(\boldsymbol{c}_j - \boldsymbol{c}_i)}{\sum_{j=1}^{N} K_h(\boldsymbol{x}_j - \boldsymbol{x}_i) K_b(\boldsymbol{c}_j - \boldsymbol{c}_i)} \tag{23}$$

where, following Machin and Gibbons (2001), $K_b(\cdot)$ is a bivariate normal kernel such that;

$$K_b(\boldsymbol{c}_j - \boldsymbol{c}_i) = \frac{1}{2\pi}\exp\left(-\frac{1}{2}(\boldsymbol{c}_j - \boldsymbol{c}_i)B^{-1}(\boldsymbol{c}_j - \boldsymbol{c}_i)\right) \tag{24}$$

and B is a 2×2 bandwidth matrix in which the diagonal elements take the values b^2 and the off diagonals are zero. Choosing the bivariate normal as the kernel function allows properties closer to property i to be more important in determining the conditional mean functions. Similarly, using a diagonal bandwidth matrix with constant diagonal elements ensures that properties equidistant from property i exert identical influence in the calculation of the conditional means.

Once again we can simplify notation by expressing the data in differences from the spatially smoothed conditional expectations; $\breve{y}_i = \ln \breve{P}_i = \ln P_i - E[\ln P \mid \boldsymbol{x}_i, \boldsymbol{c}_i]$ and $\breve{z}_i = z_i - E[z \mid \boldsymbol{x}_i]$. Stacking the $\breve{y}_i$ data to form the $N \times 1$ matrix $\breve{\boldsymbol{Y}}$ and the

$\breve{z}_i$ vectors to form the $N \times k$ matrix $\breve{Z}$ allows the parameters of the PLSS model to be estimated according to;

$$\hat{\beta}^{SS} = \left(\breve{Z}'\breve{Z}\right)^{-1}\breve{Z}'\breve{Y} \qquad (25)$$

Again, an important issue is the choice of bandwidth parameters *h* and *b*. As previously we choose *h* through cross-validation and allow for an adaptive kernel. In contrast, we impose a predetermined value on the spatial smoothing bandwidth, *b*.

This decision is motivated by a number of factors;

- In practical terms, adequate estimation of the conditional expectations in (23) can only be achieved if we ensure that the area of spatial smoothing is large enough to corral a sufficient quantity of data points.
- The primary objective of this research is to investigate the impact of noise pollution on property prices. For rail and road traffic, this is a relatively localised phenomenon; noise environments can change markedly over tens of metres and will almost certainly differ over hundreds of metres. To ensure that we can identify these localised phenomena, spatial smoothing of the regression data must operate at a somewhat larger geographical scale.[30]

A suitable scale of spatial smoothing was adjudged to be an area roughly equal to the area covered by a ward. Clearly, choosing such a scale allows an interesting comparison between the results *spatial constants model* and the *spatial smoothing model*. As a result, the spatial smoothing bandwidth, *b*, is set to a distance roughly equal to the radius of a ward[31]. The bandwidths reported in Table 2 differ across market segments reflecting differences in the size of the wards from which properties in that market segment are drawn.

---

[30] In contrast, the use of spatial smoothing (or for that matter spatial constants) will tend to reduce the ability of the model to pick out the influence on property prices of less localised spatial phenomena such as aircraft noise.

[31] In particular, we set the bandwidth for each market segment as half the average of the 20% longest distances separating properties in that market segment that are located in the same ward.

Table 2: Spatial smoothing bandwidths

Market Segment	Bandwidth (m)
1. Low Income White	1,744
2. Low Income Ethnic	1,484
3. Young FTB	1,652
4. Middle Income White	2,073
5. Middle Income Ethnic	1,738
6. High Income Large	1,981
7. High Income Small	2,274
8. High Income Medium	2,397

A priori the PLSS model is considered the better specification of the HPF. In contrast to the PLSC model, the spatial smoothing model considers information drawn from all properties in the environs of a property in adjusting for location specific variation in property prices. Further, it attaches greater weight to more proximate properties than more distant properties. Finally, including location as an argument in the nonparametric part of the PL model introduces substantially greater flexibility to the modelling of unaccounted locational factors than constraining these effects to a handful of parametric constants.

4. Spatial Correlation

Despite the use of a spatial smoothing estimator, the possibility remains that a variety of features of each property's location are not captured by the model. Since these features are held in common by properties in close proximity, it is possible that the regression residuals will exhibit spatial correlation.

As described in Section 3 of Chapter 3, the presence of spatial correlation in regression residuals ensures that OLS estimation will return inefficient estimates of the model parameters and biased estimates of the parameter's standard errors.

Over recent years, the existence of spatial autocorrelation has received a great deal of attention in the hedonic literature (e.g. Dubin, 1992; Can, 1992; Pace and Gilley, 1997; Basu and Thibodeau, 1998; Bell and Bockstael, 2000). Under the assumption that the spatial processes can be modelled as a nuisance process, researchers have tended to focus on the spatial error dependence (SED) model. In this context the SED model is given by;

$$\widetilde{\boldsymbol{Y}} = \widetilde{\boldsymbol{Z}}\boldsymbol{\beta}^0 + \boldsymbol{\varepsilon} \tag{26}$$

where

$$\boldsymbol{\varepsilon} = \rho \boldsymbol{W}\boldsymbol{\varepsilon} + \boldsymbol{u} \tag{27}$$

and $\widetilde{\boldsymbol{Y}}$, $\widetilde{\boldsymbol{Z}}$ are defined as before and can be replaced by $\breve{\boldsymbol{Y}}$, $\breve{\boldsymbol{Z}}$ for the spatial smoothing estimator. $\boldsymbol{\beta}^0$ is the $K \times 1$ vector of "true" parameters and $\boldsymbol{\varepsilon}$ is the $[N \times 1]$ vector of random error terms with mean zero. The nature of the spatial error dependence is defined by equation (27). Here $\boldsymbol{W}$ is an $N \times N$ weighting matrix, ρ is the error dependence parameter to be estimated and $\boldsymbol{u}$ is the usual $N \times 1$ vector of random error terms with expected value zero and variance-covariance matrix $\sigma^2\boldsymbol{I}$. Notice that $\rho = 0$ implies $\boldsymbol{\varepsilon} = \boldsymbol{u}$ and there is no spatial dependence in the data.

Rearranging (27) we find that;

$$\boldsymbol{\varepsilon} = (I - \rho \boldsymbol{W})^{-1}\boldsymbol{u} \tag{28}$$

which indicates that the error terms $\boldsymbol{\varepsilon}$ have a non-spherical variance-covariance matrix $\sigma^2(\boldsymbol{I} - \rho\boldsymbol{W})^{-1}(\boldsymbol{I} - \rho\boldsymbol{W}')^{-1}$. Further, the error in the SED model can be seen to be made up of two parts; a purely random element and an element containing a weighted sum of the errors on nearby properties. The association between one property and another is contained in the weighting matrix, $\boldsymbol{W}$. As in Chapter 4, the diagonal elements of the weighting matrix are zero, whilst the off-diagonal elements that represent the potential spatial dependence between observations, are non-zero only if properties are closer than some predetermined distance, d. Further we adopt a binary weights matrix in which the w_{ij}th element of $\boldsymbol{W}$ is initially set to one if the i^{th}

and j^{th} property are located within d metres of each other, otherwise that element is set to zero.

Clearly, the choice of d is important. Here we follow a similar line of logic to that used to determine the spatial smoothing bandwidth b. First we note that enumeration districts (EDs), the smallest spatial unit at which census data is available, are defined so as to isolate regions with relatively homogenous characteristics. The size of EDs, therefore, provides a guide to the spatial area over which localised similarities between properties and their inhabitants are likely to hold. As such we choose to define d for each market segment as the average radius of EDs inhabited by members of that market segment[32]. Since EDs vary in size across the cityscape, the value of d as listed in Table 3, also differs across market segments, averaging around 250m.

Table 3: Characteristics of the spatial weights matrices

Market segment	d (metres)	Characteristics of the Spatial Weights Matrix			
		Obs.	Avg Associations per Obs.	Max. Assocs	Num. with no Assocs
1. Low Income White	257	1,484	8.8	31	44
2. Low Income Ethnic	222	1,016	13.2	34	18
3. Young FTB	246	1,523	16.97	64	55
4. Middle Income White	236	1,362	5.54	22	105
5. Middle Income Ethnic	211	1,058	5.44	21	101
6. High Income Large	230	424	2.00	10	122
7. High Income Small	266	2,341	11.70	36	45
8. High Income Medium	321	1,433	10.92	47	45

Following normal procedure, $\boldsymbol{W}$ is row standardised such that each row's elements are made to sum to one. When $\boldsymbol{W}$ is row standardised, the product $\boldsymbol{W\varepsilon}$ equals

[32] In particular, we set the d for each market segment as half the average of the 20% longest distances separating properties in that market segment that are located in the same ED.

$\sum_j w_{ij}\varepsilon_j$, and has an intuitive interpretation; it is simply a vector of weighted averages of the errors of neighbouring observations. As Bell and Bockstael (2000) point out, row standardisation is undertaken to simplify estimation of the model. There is usually no underlying economic story supporting the procedure. Moreover, the spatial dependence parameter ρ estimated on a row standardised weights matrix must be interpreted with caution. In particular, ρ in this case is not directly equivalent to an autocorrelation coefficient.

The characteristics of the weights matrices constructed for each of the eight market segments are detailed in of Table 3.Even with a relatively restrictive approximately 250 metre cut-off, the majority of properties are associated with other properties in the same market segment. In market segment 2, for example, only 18 properties out of the 1,016 observations were further than 222 metres from another property in the sample. On average in this market segment, each property was located within 222m of 13 other properties in the sample, with at least one observation within 222m of 34 other properties in the sample. Notice that the number of associations in market segment 6 is somewhat lower than in the other market segments. One explanation of this observation is that properties in the affluent suburbs are more greatly dispersed than those in the other market segments.

The SED model can be estimated using maximum likelihood (ML) techniques in which the $\boldsymbol{u}$ vector is assumed to follow a multivariate normal distribution. However, for large samples this may be computationally prohibitive. Instead we follow Bell and Bockstael (2000) and use the generalised moments (GM) estimator developed by Kelejian and Prucha (1999). As Bell and Bockstael (2000) describe, whilst this estimator may not be as efficient as the ML estimator it possesses two advantages. First, the calculation of the estimator is fairly straightforward even with extremely large samples. And second, the GM estimator is consistent even when the error terms $\boldsymbol{u}$ are not normal.

The GM estimator is based on the somewhat weaker assumption that $\boldsymbol{u}$ are distributed $IID\left(0, \sigma^2\right)$. As Kelejian and Prucha (1999) show, this assumption allows us to construct the following three moment conditions;

$$E\left[\frac{1}{N}\boldsymbol{u}'\boldsymbol{u}\right] = \sigma^2$$

$$E\left[\frac{1}{N}\boldsymbol{u}'\boldsymbol{W}'\boldsymbol{W}\boldsymbol{u}\right] = \frac{\sigma^2}{N}Tr(\boldsymbol{W}'\boldsymbol{W}) \tag{29}$$

$$E\left[\frac{1}{N}\boldsymbol{u}'\boldsymbol{W}'\boldsymbol{u}\right] = 0$$

where the third equality results from the fact that the diagonal elements of $\boldsymbol{W}$ are set to zero.

Of course, the error term $\boldsymbol{u}$ is unobservable from a regression $\tilde{\boldsymbol{Y}}$ on $\tilde{\boldsymbol{Z}}$. Rather, we must rewrite the moment conditions in (29) in terms of $\boldsymbol{\varepsilon}$. Using (28) we get;

$$E\left[\frac{1}{N}\boldsymbol{\varepsilon}'(\boldsymbol{I}-\rho\boldsymbol{W})'(\boldsymbol{I}-\rho\boldsymbol{W})\boldsymbol{\varepsilon}\right] = \sigma^2$$

$$E\left[\frac{1}{N}\boldsymbol{\varepsilon}'(\boldsymbol{I}-\rho\boldsymbol{W})'\boldsymbol{W}'\boldsymbol{W}(\boldsymbol{I}-\rho\boldsymbol{W})\boldsymbol{\varepsilon}\right] = \frac{\sigma^2}{N}Tr(\boldsymbol{W}'\boldsymbol{W}) \tag{30}$$

$$E\left[\frac{1}{N}\boldsymbol{\varepsilon}'(\boldsymbol{I}-\rho\boldsymbol{W})'\boldsymbol{W}'(\boldsymbol{I}-\rho\boldsymbol{W})\boldsymbol{\varepsilon}\right] = 0$$

Under our assumptions, OLS estimation of our two models (20 & 25) will provide consistent estimates of the error terms $\boldsymbol{\varepsilon}$ and we label these $\hat{\boldsymbol{\varepsilon}}$. To simplify notation we follow Bell and Bockstael (2000) and denote $\hat{\dot{\boldsymbol{\varepsilon}}} = \boldsymbol{W}\boldsymbol{\varepsilon}$ and $\hat{\ddot{\boldsymbol{\varepsilon}}} = \boldsymbol{W}\boldsymbol{W}\boldsymbol{\varepsilon}$. Thus from (30) we can build the following three-equation system;

$$G_N\left[\rho, \rho^2, \sigma^2\right]' - g_N = v_N\left(\rho, \sigma^2\right) \tag{31}$$

where the data vectors G_N and g_N are defined as;

$$G_N = \begin{bmatrix} \frac{2}{N}\hat{\boldsymbol{\varepsilon}}'\hat{\dot{\boldsymbol{\varepsilon}}} & \frac{-1}{N}\hat{\dot{\boldsymbol{\varepsilon}}}'\hat{\dot{\boldsymbol{\varepsilon}}} & 1 \\ \frac{2}{N}\hat{\ddot{\boldsymbol{\varepsilon}}}'\hat{\dot{\boldsymbol{\varepsilon}}} & \frac{2}{N}\hat{\ddot{\boldsymbol{\varepsilon}}}'\hat{\ddot{\boldsymbol{\varepsilon}}} & \frac{1}{N}Tr(\boldsymbol{W}'\boldsymbol{W}) \\ \frac{2}{N}\left(\hat{\boldsymbol{\varepsilon}}'\hat{\ddot{\boldsymbol{\varepsilon}}} + \hat{\dot{\boldsymbol{\varepsilon}}}'\hat{\dot{\boldsymbol{\varepsilon}}}\right) & \frac{-1}{N}\hat{\dot{\boldsymbol{\varepsilon}}}'\hat{\ddot{\boldsymbol{\varepsilon}}} & 0 \end{bmatrix} \quad \text{and} \quad g_N = \begin{bmatrix} \frac{1}{N}\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}} \\ \frac{1}{N}\hat{\dot{\boldsymbol{\varepsilon}}}'\hat{\dot{\boldsymbol{\varepsilon}}} \\ \frac{1}{N}\hat{\boldsymbol{\varepsilon}}'\hat{\dot{\boldsymbol{\varepsilon}}} \end{bmatrix}$$

and $v_N(\rho,\sigma^2)$ is a 1 × 3 vector of residuals dependent on the parameters ρ and σ^2.

The system of equations in (31) can be solved by nonlinear least squares (NLS) in which the parameter estimates $\hat{\rho}$ and $\hat{\sigma}^2$ are defined as those values that minimise the sum of square residuals; $v_N(\rho,\sigma^2)' v_N(\rho,\sigma^2)$.

Armed with a consistent estimate of the spatial correlation parameter, $\hat{\rho}$, the PL models (20 & 25) can be re-estimated using feasible generalised least squares (FGLS). Accordingly, the spatial constants model can be estimated by;

$$\hat{\boldsymbol{\beta}}^{SC-SA} = \left(\tilde{\boldsymbol{Z}}'\left((I-\hat{\rho}\boldsymbol{W})'(I-\hat{\rho}\boldsymbol{W})\right)^{-1}\tilde{\boldsymbol{Z}}\right)^{-1}\tilde{\boldsymbol{Z}}'\left((I-\hat{\rho}\boldsymbol{W})'(I-\hat{\rho}\boldsymbol{W})\right)^{-1}\tilde{\boldsymbol{Y}} \qquad (32)$$

where $\hat{\boldsymbol{\beta}}^{SC-SA}$ are estimates of the parameters of a PLSC model accounting for spatial autocorrelation. Similarly, the spatial smoothing model can be estimated by;

$$\hat{\boldsymbol{\beta}}^{SS-SA} = \left(\breve{\boldsymbol{Z}}'\left((I-\hat{\rho}\boldsymbol{W})'(I-\hat{\rho}\boldsymbol{W})\right)^{-1}\breve{\boldsymbol{Z}}\right)^{-1}\breve{\boldsymbol{Z}}'\left((I-\hat{\rho}\boldsymbol{W})'(I-\hat{\rho}\boldsymbol{W})\right)^{-1}\breve{\boldsymbol{Y}} \qquad (33)$$

where $\hat{\boldsymbol{\beta}}^{SS-SA}$ are estimates of the parameters of a PLSS model accounting for spatial autocorrelation.[33]

The estimators in (32) and (33) are calculated using code written by the author in the Gauss programming language. The calculations are made feasible even in relatively large sample sizes through the use of sparse matrix commands that take advantage of the relatively large number of zero elements in the weights matrix $\boldsymbol{W}$.

[33] New residuals could be estimated using $\boldsymbol{\beta}^{SS}$ or $\boldsymbol{\beta}^{SS-SA}$ and the new solution for ρ recovered in order to iterate the FGLS estimator. However, such a procedure is not relevant for large samples and this approach is not followed here.

5. Summary of Estimation Strategy

Estimation of the parameters of the HPF for each market segment is achieved through the following steps;

1. **Construct Data Matrices**: The covariate data is divided into those variables that are to be included in the nonparametric part of the model (grouped into the matrix $\boldsymbol{X}$ with typical row $\boldsymbol{x}_i$) and those to be included in the parametric part of the model (grouped into the matrix $\boldsymbol{Z}$ with typical row z_i).

   For the PLSC model a set of dummy variables denoting location within the different electoral wards are added to the $\boldsymbol{Z}$ matrix. For the PLSS model a data matrix $\boldsymbol{C}$ with typical row $\boldsymbol{c}_i$ is constructed listing the locations of each of the properties in the data set. This data will be included in the nonparametric part of the PL model.

2. **Bandwidth Selection**: Maximum and minimum values for the bandwidth of the nonparametric kernel are selected. Grid searching across this range, the bandwidth is determined as the bandwidth value, h, that minimises the cross-validation statistic.

3. **Adaptive Kernel**: Using the bandwidth, h, the density of $\boldsymbol{x}$ for the PLSC or $(\boldsymbol{x},\boldsymbol{c})$ for the PLSS, is calculated at each observation in the data set. Subsequently, a new observation specific bandwidth h_i is calculated that is adapted to the density of the data in the region of $\boldsymbol{x}_i$ or $(\boldsymbol{x}_i,\boldsymbol{c}_i)$.

4. **Nadarya-Watson Kernel Regression**: Using Nadarya-Watson kernel regression, (13) or (23) and the bandwidth h_i, the conditional expectations of $\boldsymbol{Y}$ and $\boldsymbol{Z}$ are estimated. These expectations are stripped from the data to form the matrices $\widetilde{\boldsymbol{Y}}$ and $\widetilde{\boldsymbol{Z}}$ for the spatial constants model or $\breve{\boldsymbol{Y}}$ and $\breve{\boldsymbol{Z}}$ for the spatial smoothing model, where the new matrices represent differences from nonparametric conditional means.

5. **PL Models**: Using OLS regression (20) on the data matrices $\widetilde{\boldsymbol{Y}}$ & $\widetilde{\boldsymbol{Z}}$ or (25) on the matrices $\breve{\boldsymbol{Y}}$ & $\breve{\boldsymbol{Z}}$, the parameters $\hat{\boldsymbol{\beta}}^{SC}$ or $\hat{\boldsymbol{\beta}}^{SS}$ are estimated.

6. **Spatial Dependence Parameter**: Using the OLS regression error terms the spatial dependence parameter, ρ, is estimated using the GM estimator defined by (31).

7. **Partial Linear Models for Spatial Error Dependence**: The estimated spatial dependence parameter, $\hat{\rho}$, is used to estimate the parameters $\hat{\boldsymbol{\beta}}^{SC-SA}$ according to (32) or $\hat{\boldsymbol{\beta}}^{SS-SA}$ according to (33).

8. **Cross-Validation**: Using $\hat{\boldsymbol{\beta}}^{SC-SA}$ or $\hat{\boldsymbol{\beta}}^{SS-SA}$ the cross-validation statistic is calculated according to (16). The bandwidth h is incremented by a small amount and steps 3 to 7 are repeated across the whole range of values selected in step 2. The parameter estimates are taken as the values $\hat{\hat{\boldsymbol{\beta}}}^{SC-SA}$ or $\hat{\hat{\boldsymbol{\beta}}}^{SS-SA}$ that result from estimation with the bandwidth minimising the cross-validation statistic.

6. Identifying Under-Priced Properties

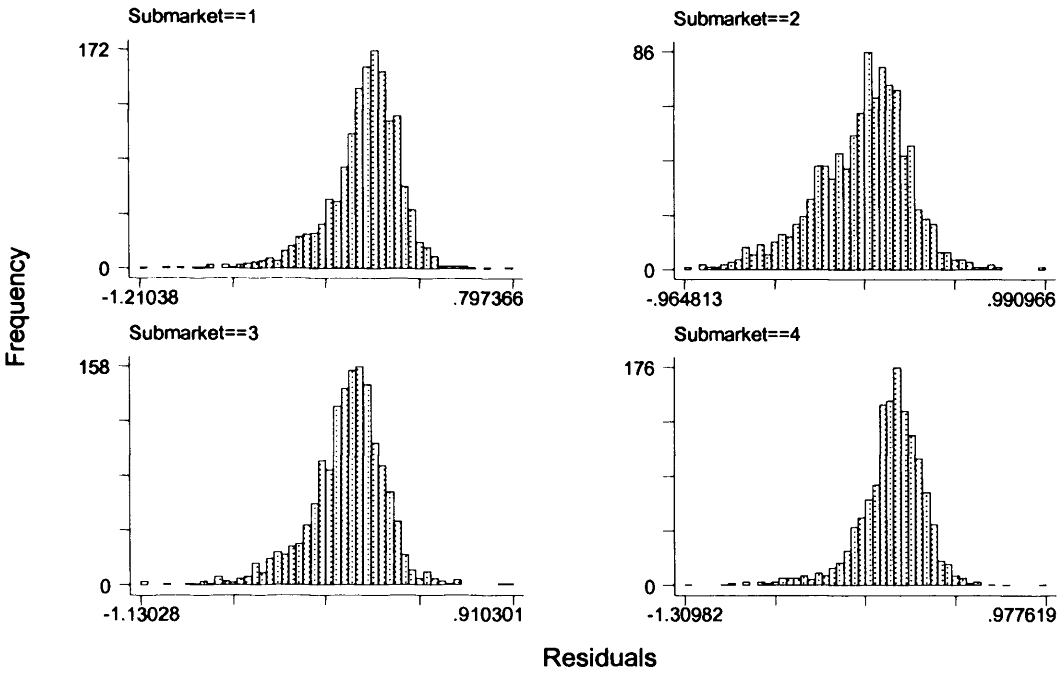
At the outset, the data was examined using simple OLS regression techniques. Regressions were run using all the covariate data whilst also including the ward level spatial constants along with linear, squared and cubed terms in the nonparametrically modelled data (properties' sale date, age, log of floor area and log of garden area). Histograms showing the distribution of residuals from these regressions are depicted in Figure 3.

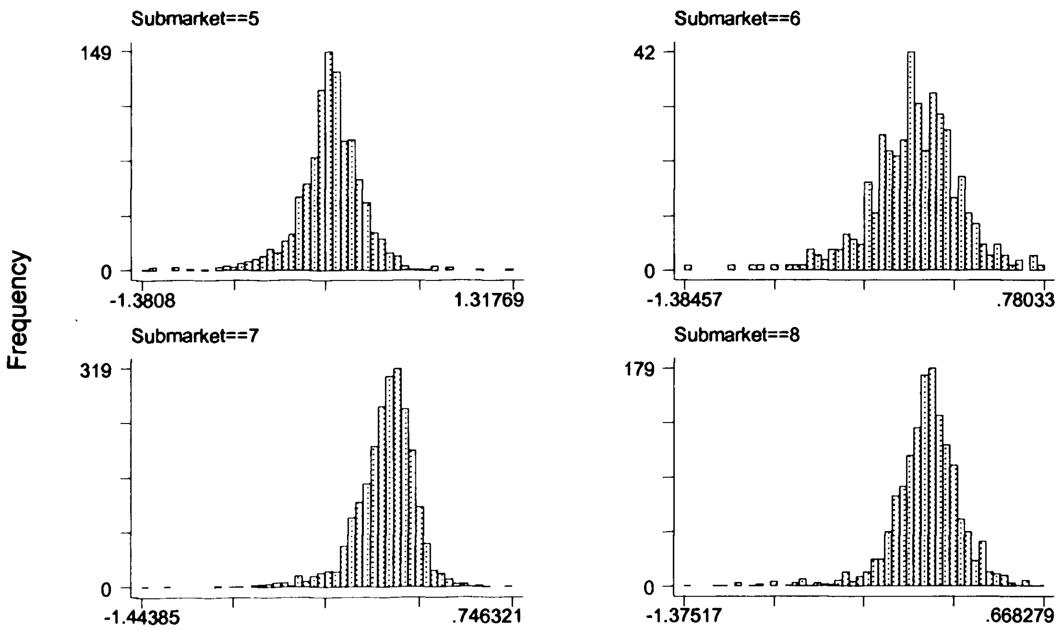
Notice that each of the distributions has a markedly long left tail. Properties with regression residuals in this tail are selling at prices well below that suggested by their known characteristics. It is unlikely that this phenomenon results from incorrect market segment classification or problems with omitted variables since there is no evidence of excessively priced properties in the right hand tails of the distributions.

By comparing properties falling in the extreme left hand tail of the error distributions with other similar properties in the same market segment it became clear that these properties were under-priced for reasons that had no basis in the data. Frequently, apparently identical houses a few doors down the road were selling for considerably greater sums. In one case 16 adjoining properties along Maypole Lane were sold within a few months of each other at prices well below the apparent market rate. Examination of recent aerial photographs of Maypole Lane provided an explanation; the houses had since been demolished to make way for a road widening scheme. The Maypole Lane properties were removed from the data set and do not feature in the results presented here.

As for the other observations it was clear that these properties, for reasons that cannot now be identified, had been sold at less than market prices. Since such transactions offer no insights into the nature of the market clearing HPF, indeed may well compromise the identification of that relationship, it was decided to remove them from the analysis. To this end, as a preliminary data cleaning exercise, all observations greater than 2½ standard deviations from the mean of the error distribution in the left hand tail were dropped. As shown in Table 4, this amounted to trimming some 2% of observations from the data.

Figure 3: Histograms of OLS residuals of HPFs by Market segment





Histograms of Residuals from OLS Regressions

Table 4: Tally of Under-Priced Observations Removed from the Data

Market Segment	Original Number of Observations	Under-priced Transactions	Final Number of Observations
1. Low Income White	1,520	36	1,484
2. Low Income Ethnic	1,034	18	1,016
3. Young FTB	1,549	26	1,523
4. Middle Income White	1,393	31	1,362
5. Middle Income Ethnic	1,074	16	1,058
6. High Income Large	430	6	424
7. High Income Small	2,388	47	2,341
8. High Income Medium	1,460	27	1,433
Total	10,848	207	10,641

7. Results: Uniqueness of Market Segment Hedonics

Full descriptions of the parameter estimates from the econometric models for each market segment are presented in Appendix C. In the rest of this Section we provide an overview of the econometric findings.

As an initial check, we test to see whether the hedonics estimated for the PLSC and PLSS models can be treated as statistically different functions. If this were not so then our contention that we have identified separate market segments through the technique of cluster analysis would be questionable.

To do this we employ a series of pairwise Wald tests to compare parameter estimates across market segments. The Wald statistic has a chi-squared distribution with k degrees of freedom, where k is the number of parameters held in common by the two models. The results of the Wald tests are presented in Tables 5 and 6. For all comparisons in both models the test statistics are significant at a greater than 95% level of confidence. The results imply that we can be reasonably certain that different price structures characterise the property markets of the different market segments.

Table 5: Wald test Chi-squared statistics for differences between property market segment hedonic price functions estimated from the PLSC model

Wald Test Statistics (p-values)

Segment	1	2	3	4	5	6	7
2	65.895 (0.000)						
3	67.199 (0.000)	89.872 (0.000)					
4	50.957 (0.001)	55.782 (0.000)	56.631 (0.000)				
5	47.613 (0.000)	67.643 (0.000)	46.611 (0.001)	46.448 (0.001)			
6	63.146 (0.002)	66.475 (0.000)	70.447 (0.000)	72.828 (0.000)	56.346 (0.000)		
7	61.179 (0.000)	65.257 (0.000)	79.375 (0.000)	46.927 (0.001)	67.333 (0.000)	79.730 (0.000)	
8	74.150 (0.000)	76.675 (0.000)	102.82 (0.000)	53.311 (0.000)	47.529 (0.000)	68.103 (0.001)	47.667 (0.000)

Table 6: Wald test Chi-squared statistics for differences between property market segment hedonic price functions estimated from the PLSS model

Wald Test Statistics (p-values)

Segment	1	2	3	4	5	6	7
2	64.736 (0.000)						
3	52.547 (0.000)	75.346 (0.000)					
4	46.838 (0.001)	61.692 (0.000)	54.757 (0.000)				
5	56.272 (0.000)	57.735 (0.000)	37.094 (0.011)	45.906 (0.001)			
6	42.616 (0.002)	63.299 (0.000)	47.833 (0.000)	41.387 (0.003)	38.652 (0.007)		
7	52.013 (0.000)	72.475 (0.000)	81.484 (0.000)	56.796 (0.000)	69.757 (0.000)	43.805 (0.002)	
8	67.354 (0.000)	78.708 (0.000)	85.746 (0.000)	37.098 (0.011)	53.314 (0.000)	44.493 (0.001)	40.069 (0.005)

8. Results: Explanatory Power of the Models

Table 7 reports unadjusted R^2 statistics for various models. The R^2 statistic measures the proportion of the total variation in the dependent variable explained by the covariates included in that model.

The first column of Table 7, presents these statistics from a series of linear OLS regressions of the log of property price against all covariates; that is, both the $\boldsymbol{X}$ and $\boldsymbol{Z}$ matrices where $\boldsymbol{Z}$ contains the ward level spatial constants. These models contain up to 94 variables. In some market segments, notably market segment 6 and market segment 8, this simple model performs remarkably well returning R^2 scores of .837 and .771 respectively. In contrast in market segments 1 and 2, the simple model is less successful with R^2 scores falling to .461 and .480 respectively.

Table 7: Explanatory power of different model specifications

Market Segment	R^2 Statistics		
	Linear Model (All Variables)	PLSC Spatial Constants	PLSS Spatial Smoothing
1. Low Income White	*.461*	*.482*	*.552*
2. Low Income Ethnic	*.480*	*.543*	*.618*
3. Young FTB	*.666*	*.718*	*.751*
4. Middle Income White	*.674*	*.719*	*.691*
5. Middle Income Ethnic	*.680*	*.757*	*.778*
6. High Income Large	*.837*	*.902*	*.886*
7. High Income Small	*.597*	*.652*	*.679*
8. High Income Medium	*.771*	*.834*	*.830*

The second column of Table 7 reports results from the PL model with spatial constants. In these models, the influence of the four variables in the $\boldsymbol{X}$ matrix are now contained in the nonparametric part of the PL model. Increasing the flexibility of the functional form by using the semiparametric estimator results in significant gains in the explanatory power of the model. For the majority of the models this is in the range of a 5% increase in the R^2 score. This result supports the contention that the semiparametric model specification significantly increases the ability of the model to describe variation in property prices.

The final column of Table 7 reports results from the PL model with spatial smoothing. Here the ward level constants used in the previous model are dropped and a nonparametric spatial smoother (operating at a similar spatial scale to the area of a ward) is introduced. In the majority of cases the PLSS model fits the data better than the PLSC model. In particular, we see a marked improvement in the R^2 statistic for market segments 1 and 2 (low-income white and low-income ethnic). Noticeably it is in these two market segments that the PLSC model performs worst, perhaps indicating considerable local variation in property prices in these market segments that is not captured by differences in the covariate data. Again, the evidence lends support to the contention that the spatial smoothing model should be preferred to the spatial constants model.

9. Results: Spatial Autocorrelation

Tables 8 and 9 present results for the detection and estimation of spatial error dependence in the two PL models. These tables report the results of two tests and the estimated spatial dependence coefficient for each of the market segments calculated using PL model residuals at the optimal (cross-validated) bandwidth. [34]

The first test statistic is Moran's I statistic (Cliff and Ord, 1972). This test is predicated on normal errors and tests the null hypothesis that there is no spatial dependence between error terms (that is, $\rho = 0$). The test statistic is asymptotically distributed as a standard normal variate. For both models, the probability of the null being true can be rejected with at least a 10% level of confidence for each of the market segments. The indications are that spatial dependence of the error exists in these models.

[34] Computational details can be found in Anselin and Hudak (1992).

Table 8: Spatial error dependence test statistics and correlation coefficient for *PLSC* model

Market Segment	Test Statistics				SAR Coef. (ρ)
	Moran's I		Kelejian-Robinson		
	Stat	Prob	Stat	Prob	
1. Low Income White	*.053*	*.000*	*163.27*	*.000*	*.133*
2. Low Income Ethnic	*.007*	*.030*	*44.95*	*.776*	*.043*
3. Young FTB	*.107*	*.000*	*528.34*	*.000*	*.301*
4. Middle Income White	*.138*	*.000*	*376.67*	*.000*	*.209*
5. Middle Income Ethnic	*.037*	*.001*	*122.18*	*.009*	*.073*
6. High Income Large	*.011*	*.000*	*88.10*	*.652*	*.100*
7. High Income Small	*.103*	*.000*	*302.58*	*.000*	*.272*
8. High Income Medium	*.090*	*.000*	*269.23*	*.000*	*.209*

The second test statistic is that proposed by Kelejian and Robinson (1992). This test is valid even with non-normal errors. The test statistic is chi-squared distributed with degrees of freedom given by the number of parameters in the model. Even with the robust test the results are reasonably conclusive with the null hypothesis of no spatial error dependence being rejected with over 99% confidence in most cases.

We conclude that spatial error dependence is an important feature of this data and thereby infer that accounting for the correlation in error terms is essential to the modelling strategy.

Table 9: Spatial error dependence test statistics and correlation coefficient for *PLSS* model

Market segment	Test Statistics				SAR Coef. (ρ)
	Moran's I		Kelejian-Robinson		
	Stat	Prob	Stat	Prob	
1. Low Income White	*.064*	*.000*	*96.38*	*.000*	*.157*
2. Low Income Ethnic	*.012*	*.088*	*37.06*	*.687*	*.074*
3. Young FTB	*.122*	*.000*	*520.45*	*.000*	*.330*
4. Middle Income White	*.163*	*.000*	*189.66*	*.000*	*.240*
5. Middle Income Ethnic	*.046*	*.011*	*103.13*	*.000*	*.084*
6. High Income Large	*.181*	*.000*	*83.46*	*.036*	*.163*
7. High Income Small	*.119*	*.000*	*204.05*	*.000*	*.298*
8. High Income Medium	*.125*	*.000*	*251.50*	*.000*	*.268*

10. Results: Selected Parameter Results

Tables 10 and 11 present a selection of parameter estimated from the PLSS and PLSC models respectively. Full listings of the model results can be found in Appendix C. Estimates are reported for each of the eight market segments. The final three rows of each table list the number of parameters in each market segment regression, *K*, the optimal bandwidth selected through cross-validation, *h*, and the number of observations in each market segment, *N*. *K* differs across models since some variables (e.g. certain ward and beacon groups constants) are not relevant to property price determination in all market segments. The dependent variable for all the models is the natural logarithm of property price.

The first two variables highlighted are taken from the nonparametric part of the PL models. Technically, the values reported are not coefficients but averages of the slope of the hedonic function estimated nonparametrically for each observation. All the same, they estimate the same quantity; the change in the dependent variable brought about by a unit change in the explanatory variable. In point of fact, since both floor area and garden area enter the regression function as logarithms, the quantities listed in the table for these variables can be interpreted as elasticities. That is, they measure the percentage increase in property prices from a one percent change in the explanatory variable.

Notice that in both models, each of the eight market segments return positive elasticities. That is, property prices are increasing in floor area and garden area. A not particularly surprising result; bigger houses are more expensive, all else equal. Notice also that the elasticities for both floor and garden area are less than one. In other words, a one percent increase in floor (garden) area increases the price of a property by less than a percentage point. Indeed for the PLSC model the estimates suggest that on average across all market segments, a 1% increase in floor area precipitates a 0.46% increase in property price. In contrast, a 1% increase in garden area results in only a 0.15% increase in property price. For both floor and garden area the elasticities tend to be larger in the more affluent market segments. A similar pattern can be seen in the PLSS model, though here the average elasticities are somewhat lower amounting to 0.33 and 0.10 for floor and garden area respectively. As we shall see shortly, this pattern appears to be a consistent point of contrast. The PLSS models return generally lower values for coefficients than the PLSC models.

A possible explanation of this pattern could arise if spatial processes tended to result in larger properties being clustered in especially desirable locations and smaller properties in especially undesirable locations. We argue that the processes by which prices evolve in the property market from the dynamic interaction of the property stock with the location decisions of households and the decisions of the public and private sectors in providing amenities, will form just such a cityscape.[35]

[35] Locations with large, well-constructed properties will tend to attract relatively more affluent residents. The wealth in such an area would encourage the provision of greater amenities by the private sector. One might expect to see more shops, more bars, more restaurants in such an area. Commonly observed correlations between affluence, educational achievement and crime suggest that such areas would also be typified by better performing schools and a safer living environment. Further, the possibility exists that greater social cohesion and organisation amongst the residents of such areas would mitigate against public authorities locating perceived disamenities (e.g. landfill

Consequently, an estimator that did not adequately account for locational effects would tend to attribute a greater importance to variables describing property dimensions than was strictly accurate. The parameter estimates would tend to compound the influence of both property size and locational desirability. The relatively larger coefficients estimated by the spatial constants model may indicate such compounding, lending support to the contention that the PLSS model is the better specification.

The next four variables reported in Tables 10 and 11 are constants describing characteristics of properties. Roughly speaking the parameter estimates on these constants relate the percentage difference in the price of a property with this characteristic compared to the base case. For example, the variable "Bedrooms 2" compares the price of properties with two bedrooms to the base case, properties with three bedrooms, all else held equal. Indeed, these estimates tend to behave as expected. The majority of coefficients are negatively signed, though only in market segments 3, 7 and 8 is the coefficient statistically significant. In these three market segments one less bedroom reduces a property's price by between 2.5% and 4.5%. The estimates are very similar for both models.

Popular perception might suggest that this variable should be more important in determining property prices. Of course, in the models presented here, the overall size of the property is controlled for by including 'floor area' as a regressor in the nonparametric part of the model. As such the parameter estimates actually record the percentage difference in the price of a property for one less bedroom, *holding total floor area constant*. Hence less bedrooms should really be interpreted as 'less but bigger bedrooms' and the relative unimportance of this variable seems more acceptable.

The "WCs 2" variable compares property prices of 'two-toileted' properties with those in the base case with one toilet. In the PLSC model only one market segment returns the expected positive and significant coefficient. The PLSS model fares somewhat better. Market segments 1, 2 and 3 each have significant positive coefficients each indicating a 3.5% price premium on properties with two WCs over

---

sites, incinerators, mental-health facilities) in that locality and perhaps even encourage the provision of public amenities (e.g. parks, recycling facilities).

those with one. Once again, the general lack of significance of this variable may be ascribed to the high correlation between the number of toilets and overall size of a property, a feature for which we have already controlled.

In contrast, the garage constant provides consistently positive and in the main highly significant parameters. Further, the parameter estimates are very similar across the two models and tend to point to the possession of a garage inflating a property's market price by some 5%. Notably the only market segment in which the parameter estimate is negative, though not significant, is the inner city market segment 2. This finding may reflect the fact that, in contrast to suburban locations, few properties in the inner city are provided with a garage (mostly having been built in the Victorian era) and that in this market segment less people have access to or requirement of their own vehicle.

Parameter estimates for the 'detached house' constant show a similar pattern. The negative and insignificant parameters for market segments 1 and 2 reflect the fact that less than 1% of properties in these market segments are detached. Amongst the more affluent market segments being detached adds between 5% and 15% to the selling price of a property.

The variables describing the socioeconomic characteristics of neighbourhoods (constructed in the factor analysis) tend to be important in explaining variation in property prices. In Tables 10 and 11 we present parameter estimates for the Wealth and Ethnicity factors. The impact of neighbourhood wealth is unequivocal. In all but one case the increasing wealth of the inhabitants of an area manifests itself in higher property prices. The parameter estimates are similar across the two models.

A more interesting pattern is revealed from the ethnicity variable. In seven of the eight market segments, the increasing presence of residents from ethnic minorities tends to decrease property prices. In contrast, the parameter on the ethnicity variable in market segment 2 (Black and Asian inner city), is positive and significant. This finding is clearer in the PLSS model where the parameter is significant at the 5% level of confidence. Evidently in market segment 2 the increasing presence of residents from ethnic minorities impacts property prices favourably. Combined these two observations tend to suggest a preference for ethnic homogeneity amongst the residents of the City of Birmingham.

For the record, Appendix C records consistent patterns for the influence of neighbourhood age composition and family composition; neighbourhoods with older residents maintain higher property prices whilst properties in neighbourhoods with more young families command lower prices.

The variable for primary schools combines distance and school quality into a single index. High scores indicate increasing quality and/or ease of access. The results here corroborate anecdotal evidence and that of recent studies (for example, Gibbons and Machin, 2002) suggesting that increasing primary school quality and proximity inflates property prices. In particular the PLSS model returns positive parameter for all of the market segments though one of these, for market segment 6, is not significant. Since market segment 6 is the high-income large property market segment this finding may simply reflect the relative lack of households with young children and/or the availability of alternative educational opportunities that reduce the perceived importance of state funded educational institutions. The PLSC model does not relate such a clear pattern. With spatial constants rather than spatial smoothing only 4 of the market segments return significant parameters for this variable.

The shops variable is constructed in the same way as the primary school variable providing an indicator of the size and proximity of local commercial centres. The patterns displayed by the parameters for this variable are more complex. The PLSS model indicates that in market segments 1 and 3 the proximity of local commercial centres significantly favours property prices whilst in the affluent suburban market segments 6, 7, and 8, proximity to shops is detrimental to property prices. In the other market segments the proximity of shops has a negligible impact on property prices. A similar, though less significant, pattern is revealed by the spatial constants model. A possible, though not entirely coherent, explanation of these results is that amongst less affluent market segments proximity to shops is considered an advantage whilst amongst more affluent suburban groups differing shopping habits reduce the attractiveness of such convenience. Indeed, amongst these latter market segments the associated disamenities of commercial centres (e.g. congestion and pollution) act so as to deflate property prices.

Table 10: Selected parameter estimates from the PLSC model

Variable	Market Segment							
	1	2	3	4	5	6	7	8
ln (Floor Area)	*0.330*†	*0.487*†	*0.375*†	*0.352*†	*0.464*†	*0.480*†	*0.588*†	*0.602*†
ln (Garden Area)	*0.067*†	*0.112*†	*0.104*†	*0.145*†	*0.233*†	*0.274*†	*0.135*†	*0.180*†
Bedrooms 2	*0.017*	*0.004*	*-0.025***	*-0.001*	*-0.030*	*-0.001*	*-0.032****	*-0.044***
WCs 2	*0.030*	*0.026*	*0.040****	*0.010*	*0.015*	*-0.024*	*-0.014*	*-0.010*
Garage	*0.055****	*-0.031*	*0.064****	*0.067****	*0.068****	*0.012*	*0.031****	*0.060****
Detached House	*0.006*	*-0.037*	*0.121****	*0.151****	*0.074****	*0.166****	*0.122****	*0.105****
Wealth	*0.121****	*0.005*	*0.047****	*0.104****	*0.102****	*0.170****	*0.055****	*0.090****
Ethnicity	*-0.094****	*0.039**	*-0.043***	*-0.090***	*-0.031***	*-0.208****	*-0.080****	*-0.106****
Primary School	*0.031*	*0.070*	*0.251****	*0.198****	*0.145***	*-0.018*	*0.079****	*0.046*
Shops	*0.017*	*-0.003*	*0.042***	*0.009*	*-0.012*	*-0.013*	*-0.018****	*-0.050****
Industry A	*2E-05**	*7E-05****	*4E-05***	*2E-05*	*5E-05***	*9E-05****	*1E-05**	*2E-05***
Landfill	*5E-05****	*4E-05*	*-3E-06*	*3E-05**	*-4E-05**	*3E-06*	*2E-05**	*2E-05*
View of Park	*-0.0001*	*0.0005*	*-0.0003*	*0.0002*	*9E-06*	*-0.0001*	*-2E-05*	*0.0001*
K	**71**	**48**	**81**	**74**	**83**	**89**	**72**	**79**
h	**.588**	**.493**	**.408**	**.447**	**.418**	**.458**	**.379**	**.366**
N	**1,488**	**1,017**	**1,527**	**1,358**	**1,058**	**424**	**2,333**	**1,432**

† Average nonparametric derivative. No significance estimated.
* Significant at 10% level of confidence
** Significant at 5% level of confidence
*** Significant at 1% level of confidence

Table 11: Selected parameter estimates from the PLSS model

Variable	Market segment							
	1	2	3	4	5	6	7	8
ln (Floor Area)	*0.287*†	*0.392*†	*0.304*†	*0.196*†	*0.347*†	*0.293*†	*0.386*†	*0.454*†
ln (Garden Area)	*0.059*†	*0.077*†	*0.070*†	*0.062*†	*0.162*†	*0.172*†	*0.089*†	*0.123*†
Bedrooms 2	*0.000*	*0.007*	*-0.028***	*-0.028*	*-0.035*	*-0.002*	*-0.030****	*-0.045***
WCs 2	*0.034**	*0.036**	*0.034***	*0.023*	*0.012*	*0.013*	*-0.011*	*-0.003*
Garage	*0.043****	*-0.027*	*0.064****	*0.068****	*0.071****	*0.031*	*0.026****	*0.056****
Detached House	*-0.046*	*-0.083*	*0.051*	*0.154****	*0.087****	*0.161****	*0.124****	*0.113****
Wealth	*0.103****	*0.001*	*0.056****	*0.129****	*0.114****	*0.161****	*0.065****	*0.084****
Ethnicity	*-0.101****	*0.048***	*-0.043***	*-0.111****	*-0.045****	*-0.147****	*-0.085****	*-0.100****
Primary School	*0.113****	*0.125***	*0.289****	*0.207****	*0.107**	*0.076*	*0.074***	*0.091**
Shops	*0.019***	*-0.014*	*0.025****	*-0.005*	*-0.007*	*-0.0033*	*-0.033****	*-0.047****
Industry A	*-1E-05*	*5E-05***	*4E-05***	*1E-05*	*2E-05*	*3E-05*	*1E-05**	*2E-05***
Landfill	*5E-05****	*-6E-05***	*-1E-05*	*4E-05***	*-4E-05***	*1E-05*	*4E-05****	*3E-05**
View of Park	*-1E-04*	*0.0007***	*-0.0001*	*0.0002*	*-6E-05*	*-0.0003*	*2E-05*	*0.0003***
K	**41**	**37**	**46**	**44**	**47**	**57**	**44**	**52**
h	**.694**	**.551**	**.574**	**.918**	**.603**	**.752**	**.547**	**.538**
N	**1484**	**1016**	**1523**	**1362**	**1058**	**424**	**2341**	**1432**

† Average nonparametric derivative. No significance estimated.
* Significant at 10% level of confidence
** Significant at 5% level of confidence
*** Significant at 1% level of confidence

The variables 'Industry A' and 'Landfill' describe the distance from a property to a locational disamenity. In general, parameters for these variables are expected to be positive indicating that property prices are greater the further one moves from the disamenity. This is certainly true of the industrial site variable. Interestingly, in this case, the PLSC model records larger and more significant parameter estimates than the PLSS model.

Results for the landfill site variable are more equivocal. Here the PLSS model returns the larger and more significant estimates. In 5 of the market segments the coefficients are, as expected, positive and 4 of these are statistically significant. However, 2 of the remaining market segments return negative and significant parameter coefficients. As we cannot support the implication that households in these market segments actually prefer to be near to a landfill site we have to conclude that other locational factors not adequately controlled for by the models are driving this result.

Finally, we consider the 'views of parks' variable. Unfortunately, this variable does not add a great deal to the analysis. Once again, the PLSS model gives the stronger result. Whilst none of the parameters estimates for 'views of parks' are significant in the PLSC model, those for market segments 2 and 8 in the PLSS model are positive and significant indicating that in these market segments at least being able to look out onto a park increases the value of property.

In conclusion, the two models tend to return plausible and pleasingly significant parameter estimates. In many cases the parameters are signed as would be expected for a particular market segment and are highly statistically significant (e.g. primary schools, ethnicity). On the other hand, a few of the estimates sit a little less comfortably with out expectations (e.g. landfill sites). Overall, the PLSS model appears to outperform the PLSC model returning a greater proportion of estimates that conform with prior expectations and are statistically significant. The improvement of the PLSS model over the PLSC model is especially apparent in market segments 1 and 2. Observe the much greater preponderance of correctly signed, significant coefficient estimates in the 2nd and 3rd columns of Table 10 when compared with the same columns of Table 11.

11. Results: Noise Pollution Parameters

Tables 12 and 13 present the parameter estimates for the noise pollution variables. Recall that these are included in the HPF in a piecewise linear fashion. That is, noise pollution is assumed to have no impact on property prices until it exceeds a threshold level of 55dB. This threshold is often taken as the "background" noise level in urban environments. Since the dependent variable is the log of property price, the coefficients represent the Noise Sensitivity Depreciation Index (NSDI). In other words, the coefficient gives the constant percentage response in property price to a one decibel absolute increase in noise pollution over 55dB.

In Tables 12 and 13, the results for each source of noise pollution (road, rail and air traffic) for each market segment are given in three columns. In each case, the first column is titled *N* and indicates the number of observations in that market segment registering a level of noise pollution (from that source) above the 55dB baseline. Clearly there are many more observations of properties exposed to road noise than to rail noise. Indeed, the relative paucity of properties exposed to rail traffic noise suggests that it will be relatively more difficult to find a statistically significant relationship between property prices and rail traffic noise pollution. The same could be said of aircraft noise which is concentrated, to a large extent, in market segments 1, 4 and 7 whose members include properties located near Birmingham International Airport. The second column for each noise source in Tables 12 and 13 provides coefficient estimates and the third column implicit prices.

In accordance with prior expectations the majority of parameter estimates on road and rail noise pollution in both models are negative. In contrast the air noise parameter estimates are far more variable. Indeed, in the PLSC model, five of the seven market segments have a positive coefficient. In mitigation, the positive and significant estimate for market segment 4 is based on a paltry sample of 10 properties exposed to air traffic noise. The PLSS model performs somewhat better, though plainly neither model is able to clearly identify a price discount for air traffic noise pollution.

Table 12: Noise pollution parameter estimates from the PLSC model

Market segment	Noise Variable								
	Road			Rail			Air		
	N	Coef	Imp Pr	N	Coef	Imp Pr	N	Coef	Imp Pr
1	*307*	*.0020*	*80.18*	*27*	*-.0097***	*-391.90*	*198*	*-.0139****	*-588.21*
2	*207*	*.0018*	*57.27*	*40*	*-.0035*	*-108.84*	*0*		
3	*523*	*-.0050****	*-229.77*	*94*	*-.0091****	*-422.13*	*17*	*.0088*	*410.29*
4	*298*	*-.0029***	*-162.07*	*42*	*-.0124****	*-691.07*	*194*	*.0032*	*176.05*
5	*271*	*-.0061****	*-318.74*	*53*	*-.0139****	*-730.68*	*10*	*.0405***	*2,125.53*
6	*168*	*-.0036*	*-484.42*	*15*	*-.0128*	*-1,748.90*	*4*	*.0311*	*4,237.00*
7	*566*	*-.0031****	*-173.37*	*60*	*-.0005*	*-26.27*	*191*	*-.0062*	*-346.49*
8	*383*	*-.0032****	*-311.96*	*48*	*-.0078***	*-766.76*	*30*	*.0033*	*327.07*

* Significant at 10% level of confidence
** Significant at 5% level of confidence
*** Significant at 1% level of confidence

Table 13: Noise pollution parameter estimates from the PLSS model

Market segment	Noise Variable								
	Road			Rail			Air		
	N	Coef	Imp Pr	N	Coef	Imp Pr	N	Coef	Imp Pr
1	*307*	*.0018*	*74.38*	*27*	*-.0084**	*-338.43*	*198*	*-.0160****	*-643.54*
2	*207*	*.0035**	*107.62*	*40*	*-.0068*	*-210.79*	*0*		
3	*523*	*-.0053****	*-245.34*	*94*	*-.0063**	*-291.38*	*17*	*-.0154*	*-741.10*
4	*298*	*-.0028***	*-156.29*	*42*	*-.0135****	*-749.91*	*194*	*.0032*	*176.62*
5	*271*	*-.0055****	*-286.31*	*53*	*-.0050*	*-260.80*	*10*	*.0339*	*1,762.41*
6	*168*	*-.0021*	*-277.70*	*15*	*-.0049*	*-662.93*	*4*	*-.0230*	*-3,109.47*
7	*566*	*-.0018***	*-100.49*	*60*	*.0001*	*5.89*	*191*	*-.0064*	*-350.40*
8	*383*	*-.0025***	*-243.63*	*48*	*-.0085***	*-831.72*	*30*	*.0033*	*-319.25*

* Significant at 10% level of confidence
** Significant at 5% level of confidence
*** Significant at 1% level of confidence

Unfortunately, air traffic noise is considerably less localised than that arising from either road or rail traffic. Indeed properties over a large area will experience very similar levels of air traffic noise. A short-coming of the modelling approach adopted in this research is that much of the influence of these wide-area spatial effects will be subsumed into the spatial constants or etiolated by spatial smoothing. Indeed, econometric specifications not reported here show that when spatial effects are not modelled, parameter estimates on air traffic noise fall into line with prior expectations. Nevertheless, the author believes that modelling wide-area spatial effects is a sacrifice worth making. In particular, considerable improvement is realised in the ability of the models to identify local-area spatial effects such as those resulting from road and rail traffic.

Focusing on the road noise variables, first observe that for both models, coefficients estimated for market segments 1 and 2 are positive. Indeed, for the PLSS model the parameter estimate for market segment 2 is just significant at the 10% level of confidence. Whilst we are happy to accept that a price premium for peaceful environments may not exist in these two low-income market segments, it seems implausible to conclude that households are actually willing to pay more for noisier properties. Clearly, some important aspect of the local environments in these two market segments is not captured by our specification of the HPF.[36]

The remaining market segments return negative road noise coefficients ranging from a value of -.0036 to -.0061 in the PLSC model and from -.0021 to -.0053 in the PLSS model. In other words, our models indicate that a one decibel increase in road traffic noise can wipe off between 0.21% and 0.61% of the selling price of a property, depending on market segment. Encouragingly, for both models, five out of the remaining six market segments have coefficients that are significant at the 95% level of confidence. Reassuringly, these parameter estimates cover the range reported from studies in other markets (see Table 3 in Chapter 3).

---

[36] With respect to market segment 2, one possibility is that within large local authority housing estates quieter environments are found towards the core of the estate whilst properties on the periphery are more likely to suffer from exposure to road traffic noise pollution. Possibly, for reasons of security and access, properties on the periphery of such estates command higher prices in the private property market than those at the core. Such a pattern might induce a positive coefficient on road noise pollution since our model does not control for these other relationships.

Turning to the rail noise coefficients, observe that on the whole the estimates are relatively larger in magnitude to those found to characterise road noise. On average the PLSC model suggest an NSDI of 0.87% whilst the PLSS model gives a somewhat more conservative estimate of 0.67%.

In all but one case, market segment 7 in the PLSS model, the rail noise coefficients are negative. Indeed, in this case the PLSC model appears to return more significant results, with three market segments recording coefficients that are significant at the 99% level of confidence and a further two market segments recording coefficients significant at the 95% level of confidence. In contrast, only four market segments in the PLSS model are statistically significant and these at a generally lower level of confidence.

To sum up, the results presented in Tables 23 to 26 are generally of the correct sign and mostly have plausible magnitudes. Overall, the PLSS would appear to outperform the PLSC model. Not only does it successfully explain more of the variation in property prices (see Table 7), but when considering the whole range of parameter estimates more frequently returns correctly signed and significant coefficients. This general conclusion does not apply to the noise parameters. Here the PLSC model tends to fare better.

12. Results: Implicit Prices for Noise

As described in Chapter 1 (Equation 4) of this thesis, the implicit price of noise (i.e. the extra that must be paid for an identical property boasting one unit less noise pollution) is given by the partial derivative of the HPF according to;

$$p_{z_k}(z_k; z_{-k}) = \frac{\partial P(z)}{\partial z_k} \qquad (34)$$

where z_k is a noise variable

Since the empirical HPF estimated here is of semi-log form, the implicit price for noise can be calculated according to the specific equation;

$$
\begin{aligned}
p_{z_k}(z_k; z_{-k}) &= \exp\left(E[\ln P \mid \boldsymbol{x}_i] + (z_i - E[z \mid \boldsymbol{x}_i])\hat{\beta}_{z_k}\right)\hat{\beta}_{z_k} \\
&= \exp\left(\ln \hat{P}_i\right)\hat{\beta}_{z_k}
\end{aligned}
\tag{35}
$$

where $\hat{\beta}_{z_k}$ is the parameter estimated on the noise variable z_k.

Notice that since the hedonic function is not linear, the implicit prices faced by each household will differ according to where they decide to locate. To provide an insight into the magnitude of these values, average implicit prices for the noise variables have been listed in Tables 12 and 13 (full details for all variables can be found in Appendix C). Note carefully that these figures are prices. They represent how much more a household would have to pay, on average, to move from their current property to an identical property that was 1dB quieter. They are not welfare measures. Those we will attempt to estimate in the final part of this thesis.

13. Conclusions

The empirical analysis described in this chapter represents the state of the art in hedonic analysis. A semiparametric approach, the *Partially Linear* (PL) Model, is used to introduce considerable flexibility into the specification of the HPF. In particular, the influence of key structural characteristics of properties is modelled nonparametrically, whilst the influence of other property characteristics are captured using a linear parametric form.

Further, unaccounted spatial effects are explicitly modelled either by introducing a set of spatial constants, the *PL Spatial Constants* (PLSC) Model, or alternatively applying a spatial smoothing procedure, the *PL Spatial Smoothing* (PLSS) Model. The residuals from these regressions have been tested and are found to exhibit evidence of spatial correlation. As a result the estimation procedure is augmented by a second stage regression that accounts for this spatial autocorrelation.

Both the PLSC and PLSS models are applied to property price data for each of the eight market segments identified using model-based clustering. Statistical tests confirm that there is little doubt that different price structures characterise the property markets of the different market segments. Furthermore, differences in the parameter estimates across the different market segments support the notion that different demand conditions characterise each segment. For example, in most market segments property prices are lower in neighbourhoods with larger numbers of inhabitants from the ethnic minorities. Interestingly, the converse is true of the ethnic minority market segment.

Overall the PLSS model appears to outperform the PLSC model returning a greater proportion of estimates that conform to prior expectations and are statistically significant. This is less true when we consider the parameter estimates on the noise pollution variables. Here we find that the PLSS model returns generally more conservative and hence less significant estimates of the influence of road and rail traffic noise on property prices.

For example, the PLSC model indicates that a 1dB increase in road noise will reduce property prices by between 0.36% and 0.61% depending on market segment. In contrast the PLSS model returns a range of between 0.21% and 0.53%.

On the whole, the rail noise estimates are relatively larger in magnitude to those found to characterise road noise. On average the PLSC model suggests a 1dB increase in rail noise will reduce property prices by 0.87% whilst the PLSS model gives a somewhat lower estimate of 0.67%.

Finally, we find little evidence of a relationship between air traffic noise and property prices. This is almost certainly a result of our model specification which subsumes the influence of wide-area spatial effects into spatial constants or etiolates their impact through spatial smoothing.

Overall, the superiority of the PLSS model leads the author to conclude that these slightly more conservative estimates should from the basis of the second stage of the hedonic analysis to be described in Chapter 9.

CHAPTER 6. NONLINEARITY IN HEDONIC PRICE EQUATIONS: AN ESTIMATION STRATEGY USING MODEL-BASED CLUSTERING

1. Introduction

The previous two chapters have reported on the application of cluster analytical techniques as a means of partitioning property market data into segments. The implicit justification for this procedure was provided by the discussion in Section 3.*i* of Chapter 3. There it was argued that markets may segment if either barriers exist that prevent households in one market segment participating in another, or if a substantial number of households hold the same inelastic demand for particular property characteristics. The partitioning of data into segments was supported by the empirical evidence. In the Birmingham case study reported in the last Chapter, the estimated HPFs were found to differ significantly across the identified segments.

This chapter presents a new and very different justification for partitioning property market data. This justification is derived directly from recent developments in the theoretical modelling of markets for differentiated goods. These models were reviewed in Sections 8 and 9 of Chapter 1 and presented in detail in Appendix A. One set of models, the *property characteristics models*, examine the nature of the market-clearing HPF when prices are determined by the characteristics of the properties themselves (e.g. Ekeland et al., 2002, 2003; Heckman et al., 2002, 2003). A second set of models, the *neighbourhood models*, examine the HPF when a property's price is determined by the characteristics of the equilibrium sets of people that choose to inhabit the neighbourhood in which that property is located (Nesheim, 2002).

Both sets of models come to one clear conclusion; under all but the most contrived of assumptions, the equilibrium HPF will be highly non-linear. Nesheim (2002) even reports that for certain parameter values, a kinked price function is required in order to attain equilibrium. Clearly, these results indicate empirical analyses imposing relatively simple functional forms are unlikely to provide good approximations to the HPF.

One response to this observation is to introduce increased flexibility into the econometric specification. Indeed, an extensive literature has developed discussing this estimation strategy (see Section 3.v of Chapter 3). However, an alternative strategy is suggested by a second finding of the theoretical models; the equilibrium market is found to be characterised by lumpy provision in attribute space.

Specifically, in the neighbourhood models, households are shown to sort themselves across the urban space such that neighbourhoods with residents showing particular combinations of characteristics may be well-represented in the equilibrium market, whilst neighbourhoods with residents exhibiting other combinations may be relatively rare. Likewise the property characteristics models can generate equilibria in which there exist clusters of properties exhibiting similar combinations of attributes, whilst properties with other combinations of attributes are sparsely represented. As Heckman et al (2003) point out, "the model is capable of generating equilibria in which there are nearly gaps in the range of products marketed".

The estimation strategy we propose here exploits this insight. In particular, using the City of Birmingham data set, we examine the attributes of properties and their neighbourhoods for evidence of the clustering suggested by the theoretical models. Following the two strands of the literature, we define two initial partitions of the data. In the first, attributes of properties are used to define clusters. In the second, the characteristics of the inhabitants of neighbourhoods are used to define clusters.

If the data confirms the existence of clusters then by definition the properties within them must lie in close proximity in attribute space and by extension close to each other on the hedonic price surface. Rather than employing increasingly more general econometric specifications to capture the nonlinearity of the equilibrium HPF, our estimation strategy is to avoid estimating the HPF over the entire attribute space. Rather, we fit separate price functions for the properties in each cluster thereby forming local approximations to the hedonic price surface over the attribute area spanned by the properties in each cluster.

Of course, if the parameters of the HPF do not differ substantially over attribute space then such an estimation strategy will be inefficient. We test this hypothesis by establishing whether there are significant differences in the parameters of the HPFs estimated for each separate cluster of properties.

Furthermore, we are interested to ascertain which of the two theoretical models forms the better approximation to the processes generating the data. Applying a non-nested test suggested by Goodman and Dubin (1990) in this context, we find that the empirical model derived from clustering by neighbourhood characteristics statistically dominates the model based on clustering by property characteristics. We conclude that the neighbourhood model would appear to better represent the processes at force in the market.

The rest of this paper is organised as follows. In section 2 we discuss model-based clustering, our approach to defining clusters in the data. Section 3 describes the data collected from the City of Birmingham in the UK that is used in this study. Section 4, reports the results of the model-based clustering. Section 5 describes the results of the econometric exercise of fitting HPFs to the different partitions of the data. Finally, Section 6 reports on the application of non-nested tests used to compare the property characteristics and neighbourhood models.

2. Identifying Clusters in Property Market Data

2.i. Segmentation or Sorting?

There is a long-established literature on the existence and identification of housing submarkets within an urban area that bears some resemblance to the work presented here (e.g. Straszheim, 1973, 1974; Ball and Kirwan, 1977; Schnare and Struyk, 1976; Sonstelie and Portney, 1980; Goodman, 1978; Michaels and Smith, 1990; Allen et al., 1995; Wolverton et al., 1999; Goodman and Thibodeau, 1998, 2003). However, contrary to the argument advanced in this Chapter, these papers motivate the existence of clusters of properties exhibiting different pricing structures through imperfections in the market mechanism. For example, Goodman and Thibodeau (2003) state that "due to either supply- or demand-related factors, the normal arbitrage that would be expected to equalize prices both within and across metropolitan areas may work either slowly, or not at all". Likewise Can (1992) states that "... varying attribute prices ... indicate the presence of independent price schedules, thus the existence of a segmented market. The presence of geographic submarkets violates the assumption of a long-run equilibrium in urban housing markets since there will be independent hedonic price schedules within a single metropolitan area reflecting the demand and supply structures of submarkets."

Of course, the theoretical literature alluded to in the introduction (and described in detail in Chapter 1) paints a quite different picture of the mechanisms at work in property markets. In particular, it shows that differences in prices across urban areas are not the result of market imperfection or disequilibrium, but rather are an integral part of the price mechanism establishing equilibrium in the property market.

Within the housing submarket literature, therefore, the definition of submarkets has tended to be dominated by the identification of property or neighbourhood characteristics that define market barriers. For instance Goodman and Thibodeau (2003) suggest that racial discrimination may produce separate submarkets for those of different ethnic origin, or that distinct sub-populations of households with strong preferences for either newly constructed properties or for historic properties may segment property markets according to the ages of properties.

In this application, however, partitioning of the data is not motivated by the supposed existence of different market segments but by the prediction of the theoretical models that property markets in equilibrium may be characterised by lumpy provision in attribute space. That is, that the market may be well-provided for certain combinations of property or neighbourhood characteristics and sparsely-provided elsewhere.

The existence of such clusters of properties is distinct from the notion of market segmentation. As such, our approach to identifying clusters is not shackled by the need to provide a formal definition of the process driving market segmentation or to formally define the property or neighbourhood characteristics by which such segments should be delineated. Rather in this paper, the data itself is used to inform on the pattern of clustering in the property market. The method by which we propose allocating properties to clusters is known as *model-based clustering*.

Clustering techniques have seen some application to the classification of properties into submarkets, notably Abraham et al. (1994), Goetzmann and Wachter (1995), Hoesli et al. (1997) and Bourassa et al. (1999). Though, since this literature is predicated by the existence of barriers to the attainment of market equilibrium, these papers do not provide a coherent justification for the use of data driven clustering techniques. Furthermore, these studies all use relatively simple clustering algorithms that provide no independent statistical indication of the nature or number of clusters to be found in the data.

2.ii. Model-Based Cluster Analysis

The basic aim of cluster analysis is to sort observations into a classification based on a set of P *clustering variables* defining the characteristics of each observation. A common starting point is to define each observation as a point in the P-space defined by these variables. Clusters are concentrations of observations falling into the same region of this P-space. Individual observations can be classified according to their proximity to different clusters.

In recent years a number of new approaches to identifying clusters in data have been proposed. As discussed in detail in Section 6 of Chapter 4, one approach that has shown particular promise is that of model-based clustering. Rather than repeat the formal presentation of Chapter 4, here we provide a brief overview of model-based clustering and describe improvements on the basic model used in this analysis that have not been presented previously.

The fundamental assumption of model-based clustering is that each data point is drawn from a population of such points constituting all the members of a cluster. Moreover, the location, size and shape of this underlying population can be approximated by a probability distribution. Assuming a Gaussian distribution, for example, would imply that clusters are ellipsoidal. It would also assume that the likelihood of observing data points belonging to a particular cluster is greater near the mean location of that cluster than at its periphery. The data observed by the researcher is the composite of data points drawn from a finite mixture of such clusters.

This model may be extended by allowing for data points that do not belong to any cluster. Banfield and Raftery (1993) suggest that such observations can be modelled as draws from a homogeneous Poisson process. That is, having removed observations belonging to clusters, the distribution of the remaining data points is one in which the expected number of "noise" observations in any location in the P-

space defined by the clustering variables is identical. The existence of "noise" observations adds an extra component to the mixture distribution. That is there is a constant $1/V$ density of observations over the P-space spanned by the clustering variables where V is the volume of that space.

Suppose we assume that there are M clusters in the data, then the mixture model describing the pattern of clustering can be formalised into the likelihood function;

$$L_M(\boldsymbol{Z} \mid \boldsymbol{\theta},\boldsymbol{\pi}) = \prod_{i=1}^{N} \left[\frac{\pi_0}{V} + \sum_{j=1}^{M} \pi_j f_j(z_i \mid \boldsymbol{\theta}_j) \right] \qquad (1)$$

where $\boldsymbol{Z}$ is the $N \times P$ matrix of data (with typical row z) by which the N observations are to be clustered; $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_M)$ is the vector of parameters associated with the assumed distributions of the M clusters, $\boldsymbol{\pi} = (\pi_0, \pi_1, \ldots, \pi_M)$ is the vector of mixing proportions, $f_j(z \mid \boldsymbol{\theta}_j)$ is the probability density function (pdf) used to model the distribution of observations belonging to cluster j; $\boldsymbol{\theta}_j$ $(j = 1, 2, \ldots, M)$ is the vector of parameters associated with the pdf of the j^{th} cluster; π_j $(j = 1, 2, \ldots, M)$ is the mixing proportion of the j^{th} cluster and π_0 is the mixing proportion of the background noise such that $\sum_{j=0}^{M} \pi_j = 1$ and $\pi_j \geq 0$ $(j = 0, 1, \ldots, M)$.

If we assume that the density of observations belonging to each cluster can be approximated by a P-dimensional Gaussian distribution, then $\boldsymbol{\theta}_j$ comprises the elements of the P-vector $\boldsymbol{\mu}_j$, which determine the mean location of the j^{th} cluster in P-space, and the distinct elements of the $P \times P$ covariance matrix $\boldsymbol{\Sigma}_j$, which determine the clusters geometric proportions.

To allow for comparison of different assumptions concerning the geometric characteristics of the different clusters, Banfield and Raftery (1993) reparameterise each covariance matrix $\boldsymbol{\Sigma}_j$ using the eigenvalue decomposition;

$$\boldsymbol{\Sigma}_j = \lambda_j \boldsymbol{D}_j \boldsymbol{A}_j \boldsymbol{D}'_j \qquad (j = 1, 2, \ldots, M) \qquad (2)$$

where $\boldsymbol{D}_j$ is the matrix of eigenvectors, λ_j is the first eigenvalue of Σ_j, and $\boldsymbol{A}_j$ is a diagonal matrix with diagonal elements $1 = \alpha_{1j} \geq \alpha_{2j} \geq \ldots \geq \alpha_{pj} > 0$.

The advantage of Banfield and Raftery's decomposition is to isolate different geometric properties of each cluster into different components. Hence λ_j determines cluster *volume*, $\boldsymbol{D}_j$ cluster *orientation* and $\boldsymbol{A}_j$ other properties of the cluster *shape*. Thus imposing the restriction $\lambda_j = \lambda \; (j = 1, 2, \ldots, M)$ enforces equality of volume across all clusters. Similarly, imposing the restriction $A_j = I \; (j = 1, 2, \ldots, M)$, where I is the P-dimensional identity matrix, generates strictly spherical clusters. Clearly differing combinations of restrictions imply different imposed similarities between clusters. One particular advantage of model-based clustering is that it provides a formal framework in which such restrictions can be compared.

For a given set of geometric restrictions and a hypothesised number of clusters, the clustering model can be simply estimated by maximising the likelihood function in (1) through application of the EM algorithm (Celeux and Govaert, 1995; Fraley and Raftery 1998). As well as providing maximum likelihood estimates of the mixing proportions, and the location and covariance structure of each cluster, a direct output of the EM algorithm is a set of $M + 1$ probabilities for each observation. The first M elements in this set provide estimates of the probabilities of an observation belonging to each of the M clusters, whilst the last element estimates the probability that the observation is part of the noise. By identifying the option with the highest probability, each observation can be allocated to either the noise or to one of the clusters. Such an allocation represents a maximum likelihood classification of the data.

A problem that remains is how to choose between clustering solutions allowing different numbers of components and differing parameterisations of cluster shapes. In contrast to other clustering algorithms, the probabilistic basis of model-based clustering provides a framework within which these comparisons can be made.

The approach followed here is that outlined in Fraley and Raftery (1998). In the first instance select a range for M, the number of clusters. Then select a series of parameterisations of the covariance matrix by applying one or more equality restrictions to (2). For each value of M and each parameterisation, the EM algorithm can be used to calculate the maximum likelihood estimates of the model parameters. Using these estimates, the Bayesian Information Criterion (BIC) for each model can be computed according to;

$$BIC_g = 2 \ln L\left(\pi_g^*, \theta_g^*\right) - \upsilon_g \log(N) \qquad (3)$$

where g indexes the particular model being evaluated, π_g^* and θ_g^* are the maximum likelihood estimates of π and θ respectively and υ_g are the total number of independent parameters in π_g^* and θ_g^*. If a BIC statistic is calculated for two different models, the difference between their BIC statistics is what will indicate the superiority of one model over the other. If the difference is large enough, one can be reasonably certain that one model gives a better fit than the other.

2.iii. Initialisation of the EM Algorithm

The EM algorithm decomposes the problem of maximising the mixture model log-likelihood into a series of relatively simple calculations. As described by Fraley and Raftery (2002a), this simplicity comes at a cost. In particular, the rate of convergence of the algorithm may be very slow and may encounter difficulties if there are a large number of clusters or the data is ill-conditioned. As with all maximisation problems, the chances of reaching a satisfactory solution are greatly enhanced by initialising the algorithm with reasonable starting values.

In this application, we follow the procedure suggested by Fraley and Raftery (1998). The data is first partitioned into those observations that are thought to fall into the clusters and those that are thought to be part of the noise. Second, using just the denoised data, observations are initially allocated to clusters using model-based hierarchical clustering (Banfield and Raftery, 1993) with an unconstrained covariance matrix. The output from this hierarchical clustering can be illustrated as a dendrogram revealing the associations between observations. To provide a categorisation of the observations into M partitions, a section can be taken through the dendrogram at the level isolating M clusters. Fraley and Raftery (1998) propose using this categorisation to initialise the EM iterations.

One shortcoming of this approach is the onerous computing requirements of hierarchical clustering methods. As described in detail in Section 3 of Chapter 5, we adopt Posse's (2001) suggested solution to this problem. Posse's approach draws on graph theoretic approaches to clustering and gathers observations together into small groups of close neighbours that would merge early in the hierarchical clustering. A hierarchical clustering of the much reduced set of observations represented by these small groups should be little different from that based on the individual data points.

We have yet to determine how data points are ascribed to the noise rather than included in the hierarchical clustering. One possibility is suggested by Fraley and Raftery (1998) who employ a nearest neighbour denoising procedure proposed by Byers and Raftery (1998). This procedure assumes that the data can be viewed as a mixture of two homogenous Poisson processes with different intensities. Observations in the clusters are drawn from the process with the greater intensity and will tend to be closer to their neighbours. Observations from the noise will be drawn from the less intense process and will be more distant from their neighbours. The procedure works by allocating an observation to the noise or the clusters according to its proximity to its neighbours.

Here we propose an alternative procedure for allocating observations to the noise based on Posse's (2001) procedure. Since this procedure gathers observations into small groups of close neighbours. Observations that are isolated from other observations in the dataset will tend to be allocated to single observations groups. We allocate these observations to the noise whilst the remaining observations are classified using hierarchical clustering. Subsequently, the data is recombined and the partitioning of the data into noise and separate clusters is used to initialise the EM algorithm for model based clustering.

2.iv. Geographic smoothing

One last issue remains to be resolved. We might expect that for properties, geographical location will play an important role in determining submarket membership. Unfortunately directly including locational variables in a cluster analysis when observations are spread reasonably homogeneously across space, tends to result in a large number of clusters that are nearly circular when spatially mapped (Fovell, 1997). As a result we follow Posse (2001) and post-process the clustering classification to take account of the spatial information.

Here we adopt a very simple rule. The six closest observations in geographical space to each observation are identified. These seven observations are examined and their classification noted. If the majority of these observations favour one classification and this differs from the classification of the target observation then the probabilities of belonging to these two different clusters are compared. Only if the target observation is less that twice as likely to belong to its current classification is the classification switched. This spatial smoothing rule is applied to all observations and the process iterated until no observations change classification.

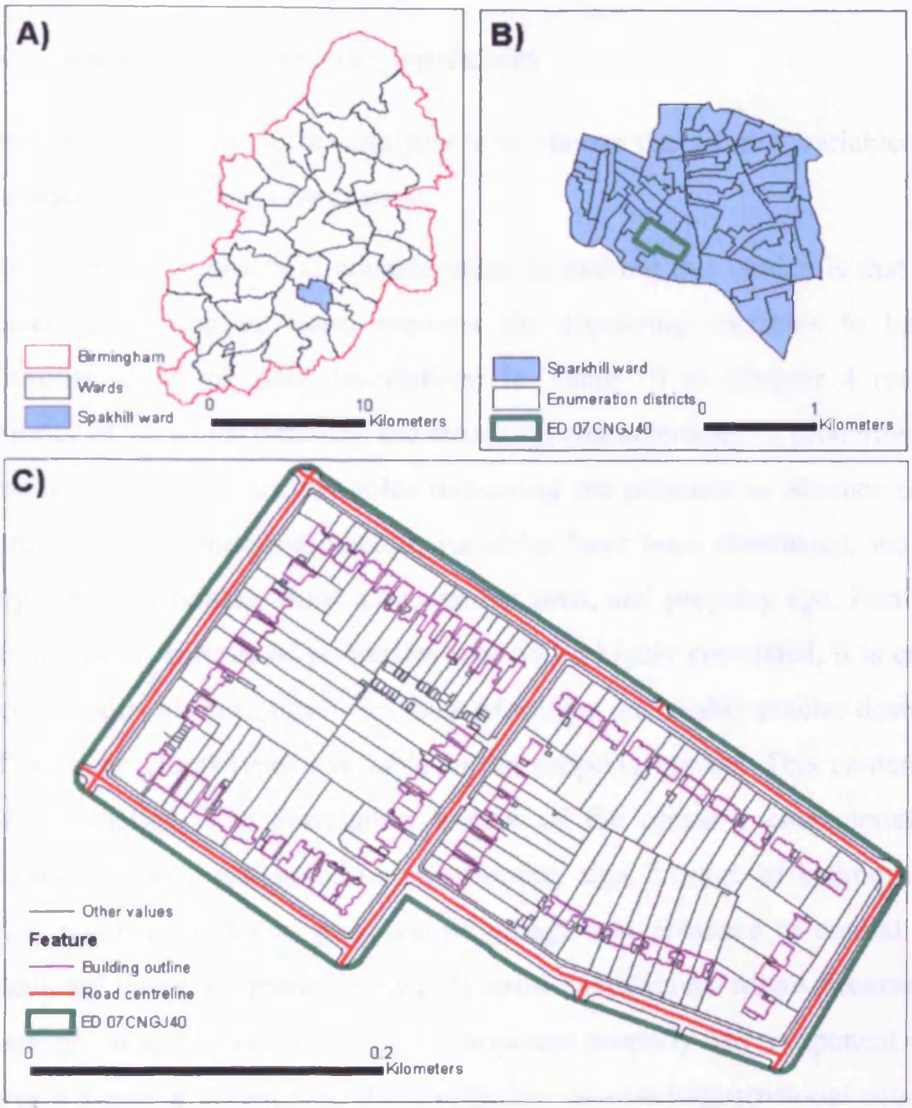
3. The City of Birmingham Dataset

The case study described in this chapter is for residential house sales in the City of Birmingham. Complete data records were successfully compiled for some 10,848 residential property transactions in 1997. Some 57 observations were excluded from the final data set for various reasons leaving 10,791 observations. Descriptions of the variables used in the hedonic analysis can be found in Chapter 4.

Of particular interest for the analysis in this Chapter, is the definition of the socio-economic characteristics of neighbourhoods. We define neighbourhoods as the smallest area over which census data is provided by the Office for National Statistics (ONS). This area is termed an enumeration district (ED). Birmingham is divided into 1,940 EDs, with each ED containing an average of 191 households. EDs are gathered into larger scale political units known as wards. Birmingham contains 39 wards such that each ward comprises an average of 50 EDs and 9,500 households. The organisation of these spatial units are shown in Figure 1.

The census provides a myriad of information on the socioeconomic characteristics of the population living in each ED. As we shall discuss in the Section 4, census data are ideal for constructing indicators of the attributes of the neighbourhood in which a property is located.

Figure 1: Hierarchy of administrative areas in Birmingham



4. Application of model-based cluster analysis to the City of Birmingham dataset

One of the key predictions of the models described in the introduction is that property markets in equilibrium may be characterised by irregular supply in attribute space. The property characteristics models predict the existence of clusters of properties exhibiting similar combinations of attributes. The neighbourhood models predict the existence of clusters of neighbourhoods inhabited by residents with similar combinations of socioeconomic characteristics. In this Section we apply the techniques of model-based clustering to the Birmingham data set in an attempt to identify clustering of these two different forms.

4.i. Choice of clustering variables

The first step in the cluster analysis is to choose the set of P variables defining the characteristics of each observation.

One exceedingly practical consideration in making this choice is that model-based clustering, as applied here, requires the clustering variables to be continuous. Examination of the data descriptions in Table 10 of Chapter 4 reveals that the majority of variables detailing the structural characteristics of properties are discrete. That is, they are binary variables indicating the presence or absence of a particular feature. Indeed, once the discrete variables have been eliminated, we are left with only three candidates; floor area, garden area, and property age. Fortunately, since the structural features of properties tend to be highly correlated, it is contended that combinations of these three variables provide a reasonably precise description of the different structural types available in the property market. This contention is borne out in practice. The correlation matrix of the property characteristics variables describing floor area, garden area, property age, number of bedrooms, number of WCs, number of floors, presence of garage and presence of central heating, was calculated using polychoric and polyserial correlations where necessary. The RV-coefficient for the floor area, garden area and property age component of this matrix gives a value of 0.790. The RV-coefficient is a multidimensional equivalent of the ordinary correlation coefficient between two variables (Robert and Escoufier, 1976)

In contrast, in defining the socioeconomic characteristics of neighbourhoods we are faced with a surfeit of candidate variables. The census data provides literally hundreds of variables describing the socioeconomic characteristics of the households inhabiting each enumerator district. As a result, we adopt a simple two-step procedure that condenses the excess of neighbourhood attributes into a more manageable set of indices or factors. In the first step, variables from the census data are grouped into five categories. These categories are as follows; variables describing the age composition of inhabitants of an ED, variables describing the family composition of households in an ED, variables describing the wealth of households in an ED, variables describing the ethnicity of inhabitants of an ED, and variables describing the education and employment of inhabitants of an ED. In the second step, the variables in each category are subjected to a factor analysis. A summary of the factor analysis is provided in Table 1.

Table 1: Factor analysis of census data describing the socioeconomic characteristics of enumerator districts

Factor Name & Description	Eigenvalues (>1)	Percent Variance Explained	Variable Loadings (>|0.50|)	
I. Household Age Composition (Using 5 variables):				
a. AGE FACTOR: Increasing Age of Inhabitants	1.76	61	% Age 18-24	-0.72
			% Age 25-34	-0.64
			% Age 50-64	0.59
			% Age > 65	0.63
II. Family Composition (Using 4 variables):				
b. FAMILY FACTOR: Increasing Proportion of Households with Children	2.66	81	% Young Family	0.86
			% Old Family	0.82
			% Age 0-10	0.78
			% Age 10-17	0.80
III. Wealth of Households (Using 4 variables):				
c. POVERTY FACTOR: Increasing Poverty of Households	3.18	97	% No car	1.00
			% Two cars	-0.85
			% Unemployed	0.85
			% Local Authority Housing	0.85
IV. Ethnicity (Using 6 variables):				
d. ASIAN FACTOR: Increasing Proportion of Asians Households	3.00	56	% Pakistani	0.96
			% Bangladeshi	0.67
			% White	-0.75
e. BLACK FACTOR: Increasing Proportion of Black Households	1.13	21	% Caribbean	0.89
			% African	0.77
V. Education and Employment (Using 15 variables):				
f. SKILLS FACTOR: Increasingly Skilled Households	2.97	34	% professional	0.61
			% diploma	0.73
			% degree	0.83

Following standard practice, for each group of variables, only factors with eigenvalues greater than one are retained. In all but one case, this results in the retention of only one factor for each category. As can be surmised from the third column of Table 1, on the whole, the retained factors capture the greater portion of the variability in the variables included in each category. The factors are rotated to aid interpretation and those variables with loadings greater than |0.50| are listed in the final column of Table 1. The loadings suggest meaningful interpretations for the dimensions captured by each factor. These interpretations are summarised in the first column of Table 1.

The final step is to define a score for each ED for each factor. In effect, EDs that exhibit high values for attributes that load positively on a factor receive high scores for that factor whilst neighbourhoods that exhibit high values for attributes that load negatively on that factor receive low scores.

The six factor scores are used as summary variables describing the major features of the socioeconomic characteristics of property neighbourhoods for use in the model-based clustering.

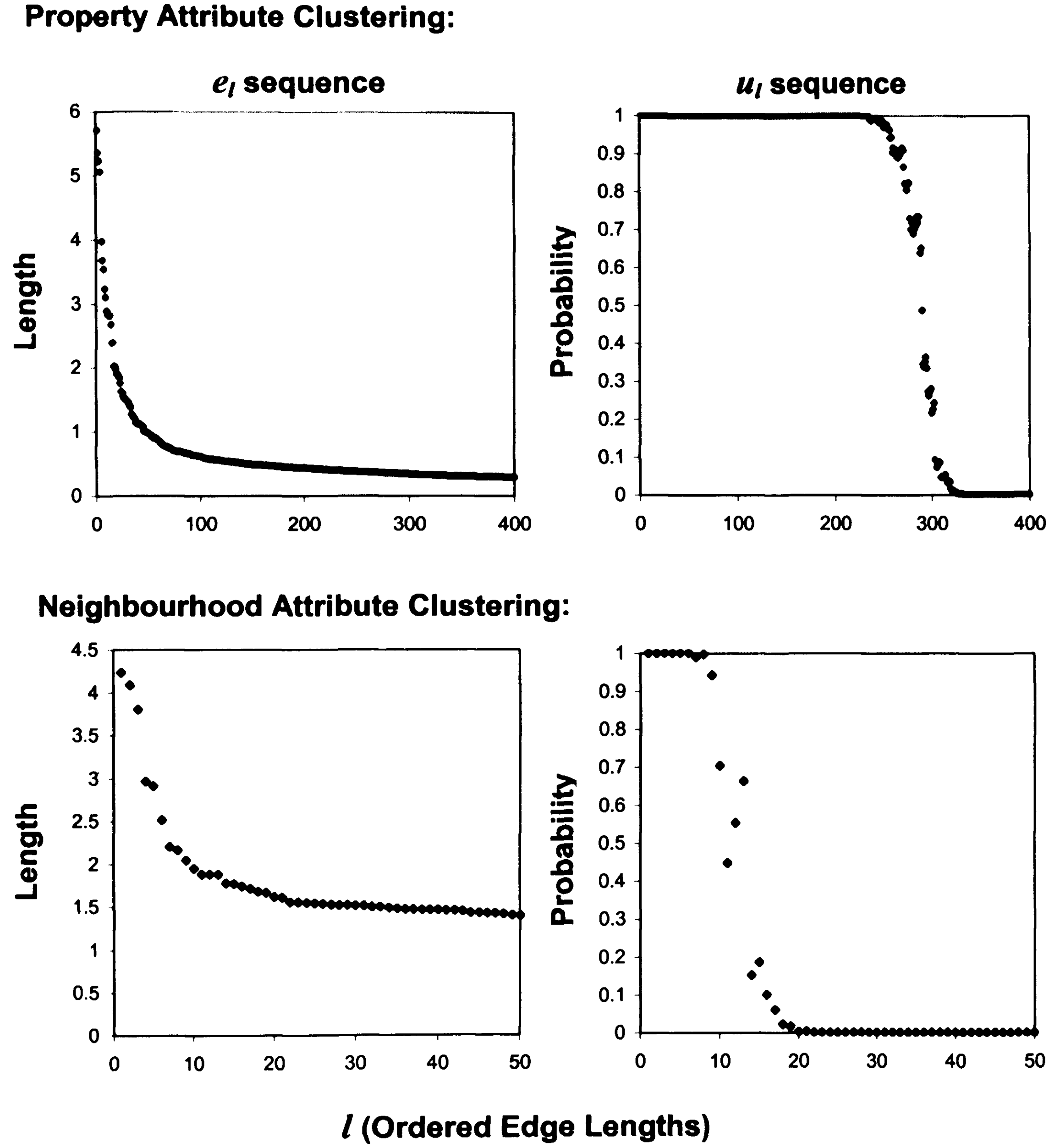
4.ii. Preliminary data partition using Posse procedure

As described in detail in Section 3 of Chapter 5, Posse (2001) proposes a procedure by which observations can be gathered together into small groups of close neighbours. He observed that members of each small cluster would merge early in a hierarchical clustering of the data. Indeed, a hierarchical clustering of the much reduced set of observations represented by these small groups should be little different from that based on the individual data points. As such, Posse's procedure greatly facilitates hierarchical clustering in large datasets where the computing requirements of working with the individual data points are extremely onerous.

In this application, we apply Posse's procedure to our two clustering datasets. The first, describing property characteristics, contains 10,791 observations and three clustering variables; floor area, garden area and property age. The second, describing neighbourhood (more precisely ED) characteristics, contains 1,940 and employs the six factors describing the socioeconomic composition of inhabitants of EDs as clustering variables.

As described in Section 6 of Chapter 4, the Posse procedure requires constructing the minimal spanning tree (MST) for each data set. The MST connects all the data points in P-space such that there is only one path connecting each pair of data points and the total length of the connections or edges joining each point is at a minimum. The edge lengths of the MST are then used to construct two sequences; the e_l-sequence and u_l-sequence. Plots of these sequences for the two clustering data sets are reproduced in Figure 2.

Figure 2: The e_l-sequence and u_l-sequence of the Minimal Spanning Tree



Posse (2001) prescribes identifying the edge length at which the u_l-sequence stabilises around zero and at which the rate of decay in the e_l-sequence drops significantly. Edges longer than this length separate observations that are more distant from each other than might be expected. As described in Chapter 3, breaking these overly long edges of the MST should leave one with groups of connected observations that are close neighbours.

From Figure 2 it is clear that the rate of decline of the e_l-sequence reduces significantly after the first 75 to 125 longest edge lengths in the case of the property attribute MST and after the first 20 to 25 longest edge lengths in the case of the neighbourhood attributes MST. Likewise, the u_l-sequence stabilises around zero shortly after the 350th longest edge length for the property attribute MST and after the 20th longest edge length for the neighbourhood attribute MST. Following Posse's (2001) proposition, therefore, we choose to peel the first 350 longest edges of the property attribute MST and the first 30 longest edges of the neighbourhood attribute MST.

As detailed in Table 2, we subsequently "prune" the MSTs so as to form a large number of roughly equal sized clusters. In the case of the property attributes data, the average number of observations in a cluster following pruning is 3.48. Similarly the average cluster size for the neighbourhood attribute is 3.41.

Furthermore, the Posse procedure isolates 765 property observations and 161 neighbourhood observations into clusters of their own. Since these singleton clusters are likely to be well-separated from other observations they are taken as an initial indication of observations that do not belong to any cluster but are part of the noise.

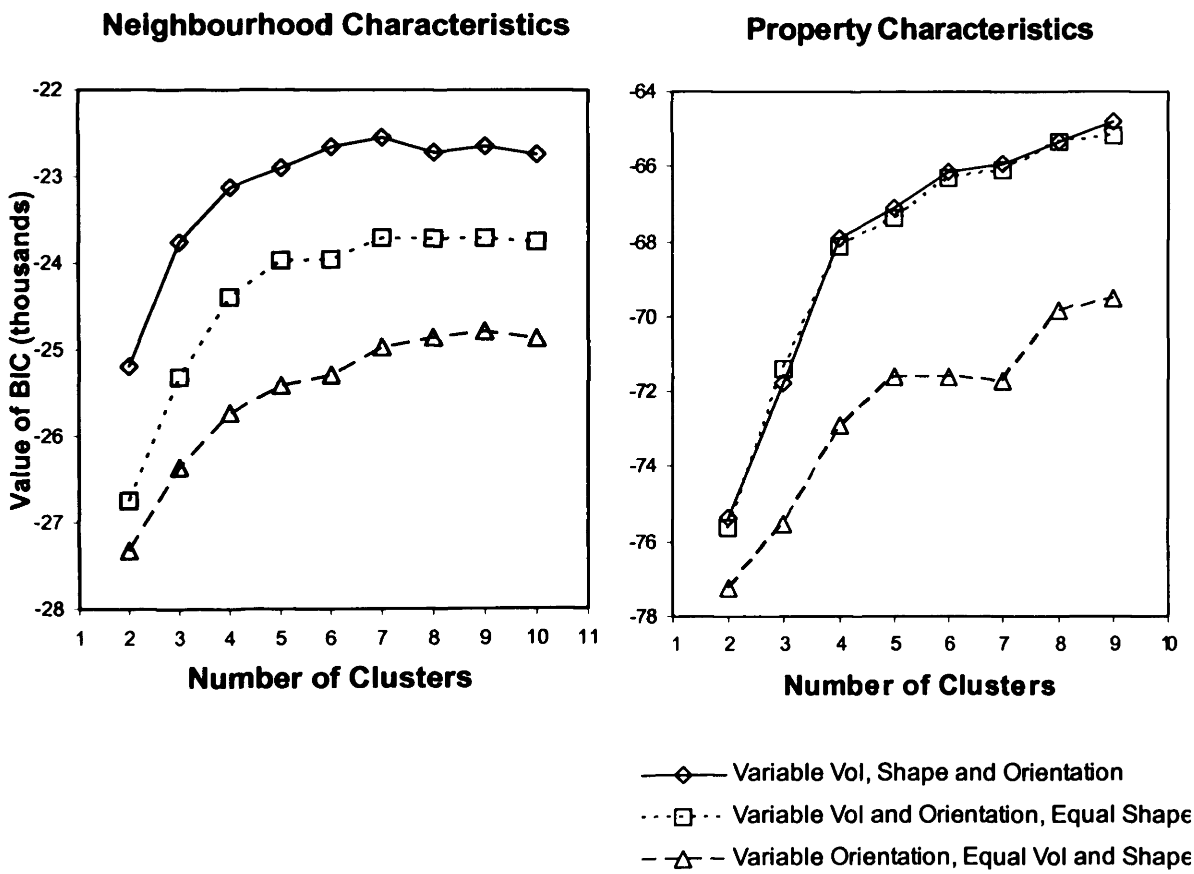
Table 2: Initial Partition of the data set using Posse's MST procedure

	Property Attribute Clustering	**Neighbourhood Attribute Clustering**
Num. Obs	10,791	1,940
Num. Peel	350	30
Num. Prune	2,500	400
Num. Clusters	3,100	569
Num. Singletons	765	161
Avg. Obs. per Cluster	3.48	3.41
Max. Obs. per Cluster	6	6

4.iii. Model based clustering with geographical smoothing

Clusters derived from Posse's procedure are used to initialise the model-based clustering algorithms. For both data sets, a variety of models corresponding to different numbers of clusters and different cross-cluster restrictions on the cluster covariance matrices have been estimated. BIC values for a selection of these models are presented in Figure 3. The three covariance models described in the figure performed significantly better than other possible parameterisations. Indeed, no other model estimated returned BIC scores that would register on these graphs.

Figure 3: *BIC* scores for clustering models assuming different numbers of clusters and different parameterisations of the covariance matrices



The BIC scores for the neighbourhood attribute clustering reveal the unconstrained model, in which different clusters may differ in size, shape and orientation, outperforms the other models. The BIC reaches a maximum at a model containing seven clusters and following Fraley and Raftery (1998) this model is selected as the one best describing the patterns of clustering in the data.

For the property attribute data the picture is less clear. A model in which the shape of each cluster is constrained to be equal performs only marginally less well than the unconstrained model. Also, there is no single maximum for the BIC scores. Rather, the BIC scores for models with progressively larger numbers of clusters tend to increase but a progressively slower rate. This pattern is not uncommon in large data sets where the BIC tends to prefer partitions with many clusters (Posse, 2001). Here we follow the suggestion of Banfield and Raftery (1993) taken up by Posse (2001) and choose the 6 cluster unconstrained model as this gives a particularly high value for the BIC in the region where the rate of change of the BIC drops significantly.

Finally, the spatial smoothing algorithm was applied to the two clustering solutions. In the case of the neighbourhood attribute partitioning the classification stabilised after 4 iterations, with some 154 EDs having changed classification. In the case the property attribute partitioning the classification stabilised after 3 iterations once 642 properties had changed classification.

Tables 3 and 4 present summary statistics that report the number of observations and the means of selected variables for each cluster. Figures 6 and 7 plot the locations of the properties in the different clusters for the two partitions of the data.

In general the neighbourhood clusters are readily interpretable. Clusters 1, 3, 4 and 5 pick out neighbourhoods that are populated, in the main, by ethnically white inhabitants. Of these Cluster 1 identifies relatively poor neighbourhoods, with low-skilled inhabitants. These neighbourhoods tend to be located to the south and west of the city but not in the city centre nor in the relatively affluent north-eastern suburbs. Cluster 4 comprises middle income neighbourhoods that are averagely skilled and relatively old. Clusters 3 and 5 pick out wealthy neighbourhoods with highly skilled inhabitants. These neighbourhoods tend to be in suburban locations with especially high concentrations in the desirable north-eastern region of the city.

In contrast, Clusters 6 and 7 define neighbourhoods whose inhabitants come mainly from the ethnic minorities. Whilst these neighbourhoods share the same inner city locations and are characterised by relative poverty and low-skilled inhabitants, they remain ethnically distinct. Cluster 6 defines neighbourhoods that are majority black, Cluster 7 neighbourhoods that are majority Asian. Perhaps unsurprisingly, average adult ages in these neighbourhoods are relatively low whilst, especially in the Asian neighbourhoods, there are a relatively large number of households with children.

Cluster 2 is somewhat more difficult to interpret. The population is ethnically diverse and comprised almost exclusively of young adults without children. Whilst the inhabitants of these neighbourhoods are relatively skilled they are only moderately wealthy. We surmise that these neighbourhoods are those inhabited by young professionals. The geographic distribution of properties in this Cluster accords with this interpretation. Neighbourhoods in Cluster 2 are located outside the inner city, but within easy commuting distance of the city centre. Further, a particularly large concentration of neighbourhoods in this cluster can be found to the south and west of the city centre, located around the University and Hospital complex.

Finally only a very few neighbourhoods cannot be assigned to one of the clusters and fall into the noise category.

Table 3: Summary of neighbourhood attribute clusters reporting the number of EDs in each cluster and the mean values for the clustering variables

Cluster	Num	Poor	Skill	Age	Family	Black	Muslim
Cluster 1	424	0.500	-0.623	0.252	0.014	-0.290	-0.409
Cluster 2	255	0.172	0.586	-0.804	-0.907	0.442	-0.297
Cluster 3	328	-1.081	0.437	0.470	-0.370	-0.602	-0.399
Cluster 4	148	0.226	-0.227	0.631	-0.259	-0.405	-0.349
Cluster 5	309	-0.990	0.875	0.392	-0.274	-0.366	-0.374
Cluster 6	256	0.915	-0.631	-0.458	0.550	1.470	0.023
Cluster 7	214	0.646	-0.538	-0.668	1.538	-0.047	2.533
Noise	6	-0.197	1.923	-1.706	-0.109	2.348	-0.142

The clusters identified by partitioning according to the age, floor space and garden size of properties are also readily interpretable.

Cluster 1 picks out modern developments. Indeed, 87% of properties in this cluster fall into Beacon Group (BG) 31 defined as standard houses built post 1953. These properties tend to be provided with gardens and cover a range of sizes and construction designs; some detached, some semi-detached and some terraced. Notice from Figure 4 that properties in this cluster are widely dispersed over the cityscape reflecting recent planning trends that have encouraged infilling rather than expansion of the urban area.

At the other extreme Cluster 2 is comprised almost exclusively of small turn-of the-century terraces with relatively small associated plots of land. 93% of these properties are classified as BG 3 or 4, that is small or medium "industrial" properties built before 1919. In accordance with the historical development of the city, these properties encircle the city centre.

Similarly, Cluster 3 identifies turn-of the-century properties located in a similar geographic region to those in Cluster 2. However, unlike Cluster 2 these are not exclusively small terraces. In fact, the properties in Cluster 3 are larger with more bedrooms and much larger gardens. Cluster 3 comprises properties constructed for the more affluent members of turn-of-the-century Birmingham society; properties that estate agents like to call "town houses" or "villas".

Table 4: Summary of property attribute clusters reporting the number of EDs in each cluster and the mean values for the clustering variables and other variates

Cluster	Num	Area	Garden	Age	Beds	%Terrace	%Detached	Price
Cluster 1	1,540	91.7	162.4	19.5	2.85	0.42	0.24	61,749
Cluster 2	2,324	95.2	84.1	93.7	2.7	0.95	0.00	38,916
Cluster 3	878	142	195.8	95	3.36	0.65	0.03	63,365
Cluster 4	1,176	97.5	266.8	49.5	2.95	0.31	0.06	57,064
Cluster 5	3,453	87.5	205.8	66	2.85	0.34	0.03	48,530
Cluster 6	1,353	136.8	506.7	57.3	3.46	0.06	0.46	107,734
Noise	67	276.3	1694.1	66.4	5.1	0.04	0.82	243,415

Clusters 4 and 5 identify standard mostly terraced or semi-detached properties with gardens. Notice in Figure 5 that the properties in Cluster 5 fall in a broad swathe that encircles the inner city. Indeed, these properties are part of the rapid expansion of Birmingham that took place between the wars. 97.5% of properties in this cluster are classified as BG 20 or 21, standard (frequently state-subsidised) properties constructed in the 1920s and 1930s. Geographically, properties in Cluster 4 appear to comprise a final ring of development surrounding the properties built between the wars. Indeed, these properties comprise standard, post-war properties. Some 70% of properties in Cluster 5 are classified as standard houses constructed between 1945-53 (BG 30) or post 1953 (BG 31).

Cluster 6 isolates the large properties in Birmingham. 86% of the properties in this cluster are detached or semidetached. They tend to have large gardens and are located in mainly suburban area with a large concentration in the desirable north-eastern region of the city.

From the descriptive statistics it would appear that many of the 67 properties allocated to the noise are the extremely large properties. The properties in the noise are mostly detached and, on average, have the most bedrooms, floor space and garden area of any of the clusters. It appears that the clustering procedure has isolated this small number of seeming outliers from the larger groupings of more moderately proportioned properties in the data.

Figure 4: Geographical distribution of properties in clusters defined by partitioning according to the socioeconomics of neighbourhoods

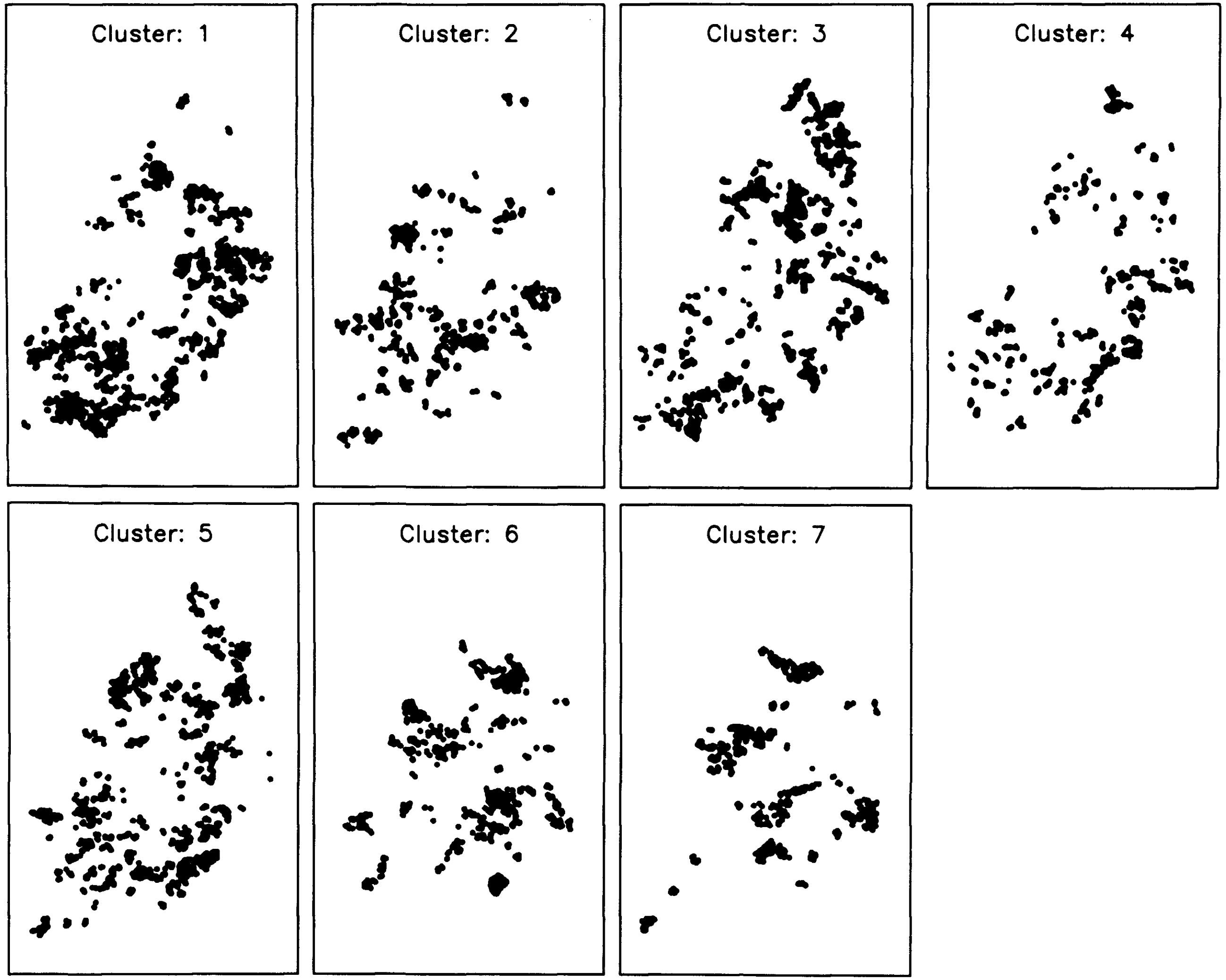
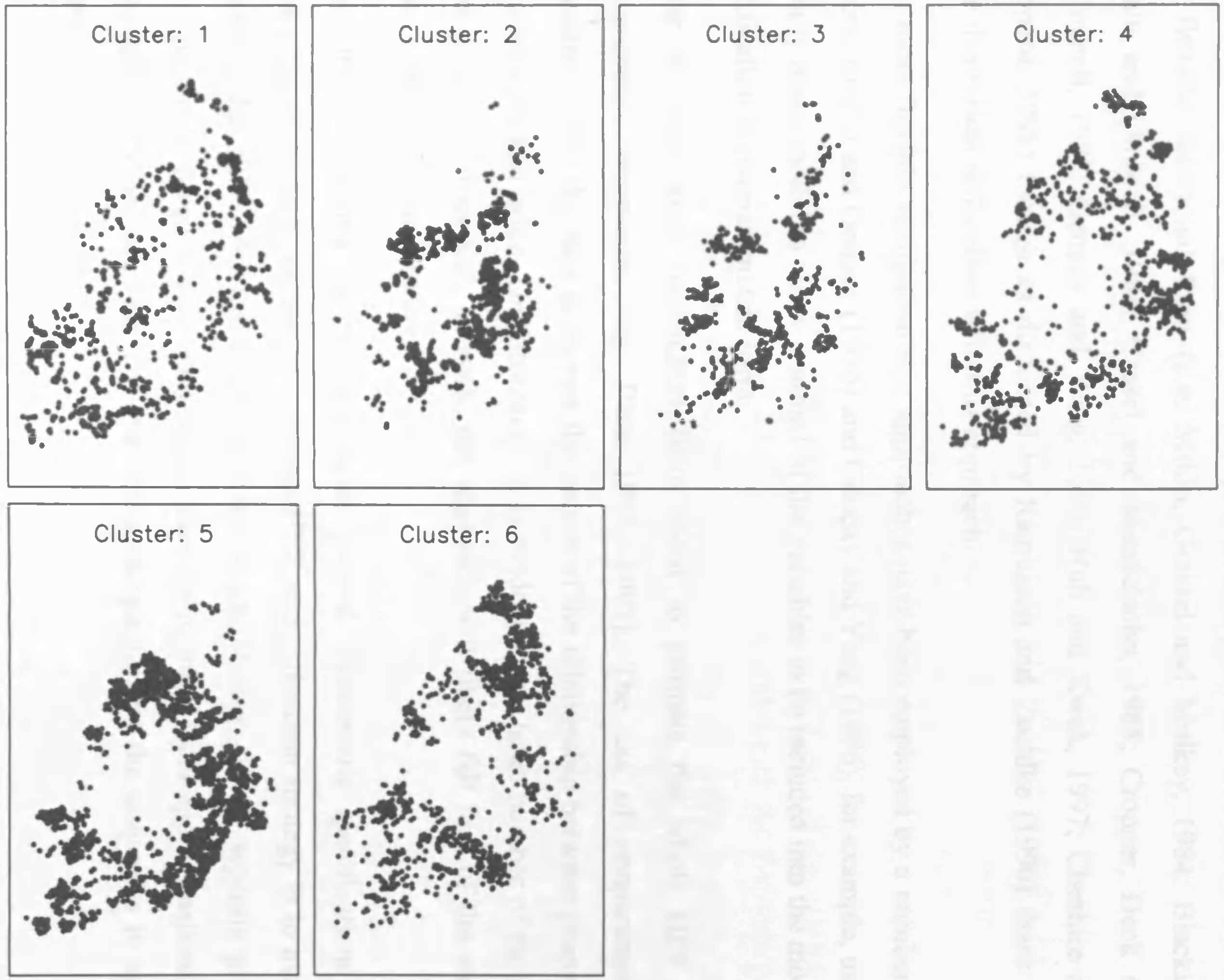


Figure 5: Geographical distribution of properties in clusters defined by partitioning according to the attributes of properties



5. Estimation of Hedonic Price Functions by Cluster

The theoretical models described in the introduction predict that the hedonic price surface may be highly non-linear. As such, following the standard procedure and fitting a simple linear regression usually with log-transformed price as the dependent variable is unlikely to provide anything but a poor approximation to the true HPF. A number of alternative estimation strategies suggest themselves.

Foremost amongst these strategies is to adopt more flexible functional forms. There is a long established literature pursuing this line of reasoning. A number of researchers have investigated the use of parametric specifications such as the Box-Cox flexible functional form (e.g. Milon, Gressel and Mulkey, 1984; Blackley, Follain and Ondrich, 1984; Cassel and Mendelsohn, 1985; Cropper, Deck and McConnell, 1988; Gençay and Yang, 1997; Huh and Kwak, 1997; Cheshire and Sheppard, 1998) though as discussed by Ramussen and Zuehlke (1990) there are some theoretical difficulties with this approach.

Even more flexible semiparametric approaches have been employed by a number of authors. Anglin and Gençay (1996) and Gençay and Yang (1996), for example, use a partially linear model to allow a subset of the variables to be included into the model specification in nonparametric form.

In the extreme, some researchers have opted to estimate the whole HPF by nonparametric regression (e.g. Pace 1993, 1995). The use of nonparametric regression allows the data to dictate the nature of the relationship between property characteristics and price. Unfortunately, it is evident that a large number of factors affect property prices and, as such, the approach will likely fall foul of the well-known curse of dimensionality.

Rather than employing increasingly more general econometric specifications to capture the nonlinearity of the equilibrium HPF, our estimation strategy is to avoid estimating the HPF over the entire attribute space. Rather, we fit separate price functions for the properties in each cluster thereby forming local approximations to the hedonic price surface over the attribute area spanned by the properties in each cluster.

For each of the two partitions of the data we adopt the following set of simple linear regression functions;

$$\ln(\boldsymbol{P}_j) = \boldsymbol{X}_j \boldsymbol{\beta}_j + \boldsymbol{e}_j \qquad j = 1, 2, \ldots, M \tag{4}$$

where j indexes clusters, $\boldsymbol{P}_j$ is the $N_j \times 1$ vector of property prices for data allocated to cluster j, $\boldsymbol{X}_j$ is the associated $N_j \times K_j$ regressor matrix, $\boldsymbol{\beta}_j$ is the $K_j \times 1$ vector of parameters and $\boldsymbol{e}_j$ is the $N_j \times 1$ vector of residuals that we assume to have $E\left[\boldsymbol{e}_j\right] = \mathbf{0}$ and $E\left[\boldsymbol{e}_j \boldsymbol{e}'_j\right] = \sigma_j^2 \boldsymbol{I}_{N_j}$. We estimate the models using ordinary least squares (OLS).

5.i. A discussion of the parameter estimates

A selection of parameter results from these two sets of linear regressions are provided in Tables 5 and 6. Full details can be found in Appendix D at the end of this thesis. For want of space, we do not discuss all the results but highlight some of the more interesting findings.

First, let us examine the partitioning based on the socioeconomic characteristics of neighbourhoods (Table 5). All in all, the parameters estimated for the structural characteristics of properties exhibit similar patterns for all seven clusters. Not surprisingly the two structural attributes describing the overall dimensions of properties, floor area and garden area, are highly significant in all clusters; the bigger the property the more it sells for, all else equal. Furthermore, in clusters where the presence of a garage and/or central-heating makes a statistically meaningful difference, it is always to make those properties more valuable.

For all clusters the parameter estimated on the age variable is negative, though it is only statistically significant at over a 90% level of confidence in two of the seven clusters. Accordingly, all else equal, properties lose market value with age. Of course, a fuller appreciation of differences in property values brought about by construction date would have to consider the parameters estimated on the eighteen Beacon Group dummy variables (to be found in Appendix D) since these also isolate important aspects of a properties age and design.

The set of dummy variables indicating the number of bedrooms possessed by a property shows a similar pattern across all clusters. Compared to the baseline case of a three-bedroom house, properties boasting more or fewer bedrooms tend to command higher prices in the market. The most statistically significant premium is for five-bedroomed properties. Of course, this is under the important caveat that all else, including floor area, is held equal.

With regards to the number of storeys over which a property is divided, there appears little to distinguish properties with one storey from the baseline case of a two-storied property. Again, to see the full picture one would need to consider the full set of dummy variables for construction type (to be found in the Appendix D) which include three variables indicating types of bungalow. In nearly all cases, and in all cases that make a statistically meaningful difference, properties with three or four stories command lower market prices than a two-storied property. It appears that the market values short, fat properties more highly than tall, skinny ones. In a similar vein, the dummy variables on construction type shown in Table 4 reveal that in all clusters, detached properties are valued more highly than semi-detached properties which are in turn valued more highly than terraced properties.

In all seven clusters, there is clear evidence of prices changing over the course of the study year (1997). Prices appear to have risen between 3% and 8% (depending on cluster) between the first and third quarter of the year, remaining stable in the final quarter.

In contrast to the structural characteristics, the influence of neighbourhood characteristics on property prices displays a number of interesting contrasts across clusters. As might be expected, property prices are depressed if the area in which the property is located is relatively poor, are inflated if the neighbourhood is inhabited by relatively highly skilled households. Perhaps not so predictably but also showing a consistent pattern across clusters we find that property prices tend to be higher in neighbourhoods with relatively older inhabitants but lower when the neighbourhood has a proportionately larger population of households with children.

In contrast, observe the parameter estimates on the Asian and Black factors. In the first five clusters, clusters whose populations appear to be majority white, neighbourhoods with larger Black or Asian contingents are characterised by lower priced properties. However, a different pattern emerges for Cluster 6, the cluster isolating neighbourhoods with mainly Black communities. Here the parameter estimates on the Asian and Black factors take the opposite sign; properties in neighbourhoods containing proportionately more Black or Asian households command significantly greater market prices. A similar pattern can be seen in Cluster 7, the cluster isolating majority Asian neighbourhoods. In this cluster, properties in neighbourhoods with proportionately more Asian households are significantly more expensive. Without wishing to over-elaborate the significance of this result, the implication is that within clusters properties in racially homogeneous neighbourhoods tend to be more highly valued than those in ethnically diverse neighbourhoods.

Consider now the locational characteristics of properties with respect to their proximity to amenities and disamenities. The parameters on the proximity to the city centre present a somewhat confused pattern, being negative and significant for some clusters, positive and significant for others. For example, proximity to the city centre deflates property prices in clusters 1 and 6 (the poor ethnically white and ethnically black clusters respectively) whilst inflating prices in clusters 3 and 7 (the wealthy and Asian clusters respectively). Since, proximity to the city centre does not induce a coherent influence on property prices across all clusters it seems likely that this variable is proxying for other features of the urban geography that are not captured by the model.

The patterns displayed by the parameters on the shops variable, which provides an indication of the size and proximity of local commercial centres, are again somewhat complex. The model indicates that in clusters 1 and 6 the (the poor ethnically white and ethnically black clusters respectively) property prices increase with proximity to shops though in clusters 3 and 7 (the wealthy and Asian clusters respectively) prices are reduced by proximity to shops. A possible, though not entirely coherent, explanation of these results is that within the less affluent socioeconomic clusters, proximity to shops is considered an advantage whilst amongst more affluent suburban groups differing shopping habits reduce the attractiveness of such convenience.

Table 5: A selection of parameters from hedonic price equations for clusters defined by partitioning according to the socioeconomics of neighbourhoods

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6	Cluster 7
Constant	8.9067***	8.2327***	8.7047***	8.4688***	8.3764***	8.4239***	8.9403***
Structural Characteristics:							
Floor Area (log)	0.3827***	0.3991***	0.4383***	0.3612***	0.4879***	0.3864***	0.3670***
Garden Area (log)	0.0838***	0.1662***	0.0973***	0.1005***	0.0940***	0.1393***	0.1446***
Garage	0.0448***	0.0579***	0.0550***	0.0350*	0.0524***	0.0607***	0.0369
Central Heating	0.0464*	0.0653*	0.0577**	-0.0283	0.1032***	0.0828**	-0.0716
Age	-0.0148**	-0.0067	-0.0096	-0.0058	-0.0204***	-0.0091	-0.0106
WCs							
One	*b*	*b*	*b*	*b*	*b*	*b*	*b*
Two	0.0243*	-0.0399**	0.0315**	-0.0059	0.0297**	-0.022	-0.0244
Three	0.0198	0.2056**	-0.0112	-0.022	0.1304***	0.0222	0.0075
Four	.	0.8627***	-0.2295*	.	0.4666**	.	.
Bedrooms							
One	0.0727	0.0675	0.0414	0.2473**	0.0351	0.3394	0.1957
Two	0.007	-0.0013	0.0127	-0.0299	0.0152	-0.0062	0.0560**
Three	*b*	*b*	*b*	*b*	*b*	*b*	*b*
Four	0.0278	0.0029	0.0165	0.0279	0.0407*	0.0677**	0.047
Five	0.0452	-0.0609	0.1349***	0.1474**	0.1459***	0.1758***	0.044
Storeys							
One	-0.07	-0.4751	-0.037	0.0449	0.1903***	0.2255	0.1281
Two	*b*	*b*	*b*	*b*	*b*	*b*	*b*
Three	-0.0481	-0.2195***	-0.1069***	-0.0672	-0.1115***	-0.0166	0.0183
Four	-0.2106*	-0.8875***	-0.4576***	-0.1522	-0.1909*	-0.1956	-0.4995**
Construction Type							
Detached	0.1396***	0.1477***	0.1220***	0.1386***	0.1087***	0.0721***	0.0884**
Semi-Detached	*b*	*b*	*b*	*b*	*b*	*b*	*b*
End Terrace	-0.0887***	-0.0981***	-0.0440**	-0.0493*	-0.0780***	-0.0309	-0.1012***
Terrace	-0.0795***	-0.0418*	-0.0647***	-0.0407*	-0.0917***	-0.0763***	-0.0833***
Sale Date							
1st Quarter	-0.0508***	-0.0313*	-0.0564***	-0.0400**	-0.0407***	-0.0716***	-0.0675***
2nd Quarter	-0.0231*	-0.017	-0.0224*	-0.0495***	-0.009	0.0246	-0.0262
3rd Quarter	*b*	*b*	*b*	*b*	*b*	*b*	*b*
4th Quarter	-0.0039	-0.0094	0.0023	-0.0171	0.015	0.0269	-0.0101

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6	Cluster 7
Neighbourhood Characteristics							
Poverty Factor	-0.0871***	-0.0736***	-0.0648***	-0.1284***	-0.0471***	-0.1260***	-0.0483**
Skills Factor	0.0628***	0.0305***	0.0524***	0.0401**	0.0662***	0.0055	0.0404**
Age Factor	0.0209***	0.0303**	0.014	0.0264*	0.0098	0.0587***	0.0404**
Family Factor	-0.0075	-0.0489***	-0.0205*	-0.0255	-0.0106	-0.0248	-0.0506***
Asian Factor	0.0137	-0.0482***	-0.0368***	-0.0697**	-0.0118	0.0314**	0.0438**
Black Factor	-0.0254**	0.0046	-0.0517***	-0.0314	-0.0520***	0.0269**	-0.011
Locational Characteristics							
City Centre	0.0001**	-0.0001	-0.0001**	-0.0001	0	0.0001*	-0.0001*
Shops	0.0230***	-0.0134	-0.0347***	-0.0023	0.0126	0.0271**	-0.0363***
Primary Schools	0.0961**	0.1593***	0.1089***	0.1766***	0.0942**	0.0134	0.0404
Rail Station	0	0	0	0.0001***	0.0000**	0	0
Park	0	0	0.0000**	0	0	0	0
Airport	-0.0001***	-0.0001	0	0	-0.0001***	-0.0001**	-0.0001
A-Type Industry	0	0.0001***	0	0.0000**	0.0000**	0	0
B-Type Industry	0	-0.0001***	0	-0.0001*	0	-0.0001***	0
Land Fill sites	0	0.0000**	0.0000**	0.0001***	0.0000*	0.0000**	0
Environmental Characteristics							
Views of Water	0.0055**	-0.0001	0	0.0029	-0.0009	-0.0008	0.0002
Views of Parkland	0	-0.0002	-0.0002	0	0.0002	0.0003	0
Road Traffic Noise	-0.0004	-0.0002	-0.0024**	-0.0037**	-0.0038***	-0.0035**	-0.0035*
Rail Traffic Noise	-0.0026	-0.0126*	-0.0086**	-0.0089**	-0.0023	-0.0046	-0.0119**
Aircraft Noise	-0.0906*	-0.1413	0.0102	-0.0637	.	.	-0.0109
K	96	90	96	93	97	85	82
N	2261	1258	2173	895	2018	1207	970
R^2	0.721	0.830	0.800	0.807	0.790	0.847	0.829
s^2	0.0455	0.0471	0.0456	0.0382	0.0457	0.0514	0.0588

b Base case for a set if dummy variables
* Significant at 10% level of confidence
** Significant at 5% level of confidence
*** Significant at 1% level of confidence

The variable for primary schools combines distance and school quality into a single index. High scores indicate increasing quality and/or ease of access. The results here corroborate anecdotal evidence and that of recent studies (for example, Gibbons and Machin, 2002) suggesting that increasing primary school quality and proximity inflates property prices. Whilst the parameters for all clusters are positive, those in the ethnic minority socioeconomic clusters (clusters 6 and 7) are not significant.

Some fairly general patterns emerge with regards to the other locational variables. Proximity to railway stations and parks tend to have little influence on property prices. In the clusters where these locational characteristics make a difference, they act so as to decrease property prices with increasing proximity. Whilst, these locational features could nominally be considered as amenities, it appears that other issues, perhaps including security and noisy activity, may detract from the benefits of proximity to either a railway station or park. Where significant, proximity to Type-A industrial processes and proximity to landfill sites act so as to reduce property prices. In contrast, proximity to the airport and proximity to Type-B industrial processes act so as to increase prices.

In accordance with prior expectations all the parameter estimates on road and rail noise pollution are negative and in the majority of cases are statistically significant. A similar pattern emerges for estimates of the parameter on the aircraft noise pollution variable, though here only one parameter estimate is statistically significant and another is positive (though not significant at a 90% level of confidence). Unfortunately, air traffic noise is considerably less localised than that arising from either road or rail traffic. Indeed properties over a large area will experience very similar levels of air traffic noise. A short-coming of the modelling approach adopted in this research is that much of the influence of these wide-area spatial effects will be subsumed into the locational constants indicating ward membership (parameter estimates for these ward constants can be found in Appendix D).

Parameter estimates for the model based on partitioning the data according to the attributes of the properties are displayed in Table 6. Conclusions concerning the impact of structural attributes on property prices for this partitioning of the data are broadly similar to those for the partitioning based on the socioeconomic composition of neighbourhoods.

In Table 6, parameters for the socioeconomic variables describe an almost identical pattern to that found with the socioeconomic partitioning of the data. Property prices are depressed if the area in which the property is located is relatively poor, are inflated if the neighbourhood is inhabited by relatively highly skilled households, tend to be higher in neighbourhoods with relatively older inhabitants but lower when the neighbourhood has a proportionately larger population of households with children. In all but cluster 2, prices tend to be driven down in neighbourhoods with higher proportions of Asian and/or Black households. Cluster 2, separates the inner city properties where the majority of Asian and Black residents of Birmingham are located. Within this cluster properties in neighbourhoods with proportionately more Asian of Black households are significantly more expensive. Again the data suggests that the market rewards ethnic homogeneity.

With regards to locational characteristics, our conclusions concerning the impact on property prices from the proximity of (dis)amenities are little changed from those arrived at for the partitioning of the data according to the socioeconomic composition of neighbourhoods. One point of contrast concerns the variable describing the proximity and quality of primary schools. Notice that the parameter on primary schools is significant in cluster 2, the cluster which isolates the inner city properties. This contrasts with the results for the socioeconomic partitioning where properties in this cluster were divided between the Asian and Black socioeconomic clusters (clusters 6 and 7 of the socioeconomic partitioning) and were found not to be significant. In contrast, we find that the only cluster in which primary school proximity and quality does not exert a significant influence on property price is in cluster 6. Since this cluster identifies the large properties in the Birmingham property market this finding may simply reflect the relative lack of households with young children and/or the availability of alternative educational opportunities that reduce the perceived importance of state funded educational institutions.

Once again the parameters on road and rail noise pollution variables are negative for all clusters. However, as a general observation, these tend to show less significance than was exhibited in the socioeconomic partitioning.

Table 6: A selection of parameters from hedonic price equations for clusters defined by partitioning according to the attributes of properties

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6
Constant	8.9319***	8.6194***	8.4596***	8.0269***	8.2942***	8.5659***
Structural Characteristics:						
Floor Area (log)	0.2795***	0.4408***	0.3860***	0.5532***	0.4595***	0.4090***
Garden Area (log)	0.0822***	0.0659***	0.1256***	0.0766***	0.0587***	0.1667***
Garage	0.0762***	0.0063	0.0349	0.0567***	0.0416***	0.0467**
Central Heating	0.0185	0.0204	-0.008	0.0701**	0.0742***	0.2206***
Age	-0.0556***	-0.0131**	0.0149	-0.0116	-0.01	-0.0281***
WCs						
One	*b*	*b*	*b*	*b*	*b*	*b*
Two	-0.0216	0.0148	0.0268	0.0211	0.0037	-0.0027
Three	0.0207	.	0.1565	.	-0.032	0.0373
Four	-0.0882	.	.	.	.	0.3829
Bedrooms						
One	-0.0434	0.1536	0.8127***	-0.0036	0.104	-0.4211***
Two	-0.0117	-0.0006	-0.0607*	0.0131	0.006	0.0404
Three	*b*	*b*	*b*	*b*	*b*	*b*
Four	0.0552**	0.0891**	0.0114	0.0323	-0.0143	0.0224
Five	0.2768***	0.2126	0.0113	.	0.0546	0.0523*
Storeys						
One	0.1283	.	-0.3993	0.0474	-0.0201	-0.0447
Two	*b*	*b*	*b*	*b*	*b*	*b*
Three	-0.1098**	0.0676	-0.1193***	-0.0598	-0.0684***	-0.1682***
Four	-0.3953***	.	-0.4555***	.	.	-0.0463
Construction Type						
Detached	0.1697***	-0.2648***	0.1527**	0.0836***	0.0611***	0.1185***
Semi-Detached	*b*	*b*	*b*	*b*	*b*	*b*
End Terrace	-0.0607***	-0.0531*	-0.036	-0.0818***	-0.0474***	-0.1310***
Terrace	-0.0760***	-0.0740***	0.0019	-0.0764***	-0.0663***	-0.017
Sale Date:						
1st Quarter	-0.0435***	-0.0756***	-0.0626**	-0.0447***	-0.0363***	-0.0555***
2nd Quarter	-0.0147	-0.0276**	0.0142	-0.0222	-0.0097	-0.0404**
3rd Quarter	*b*	*b*	*b*	*b*	*b*	*b*
4th Quarter	-0.0016	-0.0038	0.0283	0.0241	0.0039	-0.0237

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6
Neighbourhood Characteristics						
Poverty Factor	-0.1069***	-0.0496***	-0.0719***	-0.1023***	-0.0756***	-0.0163
Skills Factor	0.0078	0.0469***	0.0742***	0.0261**	0.0293***	0.0582***
Age Factor	0.0173**	0.0261**	0.0385**	0.0126	0.0370***	0.0522***
Family Factor	-0.011	-0.0386***	-0.0074	-0.0035	-0.0051	-0.0084
Asian Factor	-0.011	0.0326***	-0.0441**	0.0316	0.0072	-0.0681**
Black Factor	-0.0500***	0.0190**	-0.0346**	-0.0217	-0.0431***	-0.0632***
Locational Characteristics						
City Centre	0	0.0001	-0.0002	0	0	-0.0001**
Shops	0.0048	0.0115*	0.0117	-0.0168	0.0105*	-0.0350***
Primary Schools	0.1033**	0.0726*	0.2103**	0.1804***	0.0897***	0.0252
Rail Station	0.0000**	0.0000**	0	0.0000**	0.0000***	0.0000**
Park	0	0	0.0001*	0.0000*	0	0
Airport	-0.0001**	0	-0.0001	-0.0001***	-0.0001***	-0.0001**
A-Type Industry	0	0.0000***	0.0001**	0	0.0000**	0.0000***
B-Type Industry	0	0	0	0	0.0000***	0
Land Fill sites	0.0000**	0	-0.0001*	0.0001***	0	0.0000***
Environmental Characteristics						
Views of Water	-0.0023	0	0.0003	-0.0029	0.001	0.0004
Views of Parkland	-0.0002	0	0.0001	-0.0001	0	0
Road Traffic Noise	-0.0024	-0.0022**	-0.0052***	-0.0019	-0.0016*	-0.0017
Rail Traffic Noise	-0.0063*	-0.0074**	-0.0055	-0.0128**	-0.0042	-0.0039
Aircraft Noise	0.0092	.	.	-0.0072	-0.0123	-0.0095
K	88	80	91	87	86	101
N	1540	2324	878	1176	3453	1353
R^2	0.760	0.646	0.655	0.686	0.574	0.763
s^2	0.0387	0.0525	0.0772	0.038	0.0355	0.0486

b Base case for a set if dummy variables
* Significant at 10% level of confidence
** Significant at 5% level of confidence
*** Significant at 1% level of confidence

5.ii. *Comparison of the hedonic price functions across clusters*

The estimation strategy followed in this paper is to capture the nonlinearity of the equilibrium HPF by fitting separate price functions for the properties in each cluster. Clearly, a question we would like to answer is whether this estimation strategy makes a difference. In particular, we need to test whether the HPFs estimated for the different clusters differ from each other in statistically meaningful ways.

We do this by carrying out a series of pairwise comparisons. For example, we may wish to test the hypothesis that the parameters of the HPF estimated from the first cluster do not differ significantly from those estimated from the second cluster. That is, we wish to test the hypothesis that $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$ in the two linear regressions;

$$\ln(\boldsymbol{P}_j) = \boldsymbol{X}_j\boldsymbol{\beta}_j + \boldsymbol{e}_j \qquad\qquad j = 1, 2 \tag{5}$$

where $\boldsymbol{P}_j$, $\boldsymbol{X}_j$, $\boldsymbol{\beta}_j$ and $\boldsymbol{e}_j$ are defined as before but we also assume that $\boldsymbol{e}_j$ follows a multivariate normal distribution with mean zero and covariance matrix $\sigma_j^2\boldsymbol{I}$.

In the special case in which we can assume that $\sigma^2_1 = \sigma^2_2$ the stability of the parameters can be tested using a small sample test such as the Chow Test. This approach has been adopted by a number of previous authors in this field (e.g. Michaels and Smith, 1990; Allen et al., 1995). A quick glance across the values for s^2 (the OLS estimates of σ^2) in Tables 5 and 6 indicates that the equality of error variances is unlikely to hold true in this case. Unfortunately, when $\sigma^2_1 \neq \sigma^2_2$ the Chow test is invalid (Toyoda, 1974).

An alternative test is offered by the Wald statistic given by;

$$W = (\boldsymbol{b}_1 - \boldsymbol{b}_2)'\left(s^2_1\boldsymbol{\Sigma}_1 + s^2_2\boldsymbol{\Sigma}_2\right)^{-1}(\boldsymbol{b}_1 - \boldsymbol{b}_2) \tag{6}$$

where $\boldsymbol{b}_j$ and s_j^2 are the least squares estimates of $\boldsymbol{\beta}_j$ and σ_j^2 respectively and $\boldsymbol{\Sigma}_j$ is $\left(\boldsymbol{X}_j'\boldsymbol{X}_j'\right)^{-1}$. Nominally, this statistic has a chi-squared distribution with k degrees of freedom, where k is the number of parameters in common between the two models.[37] In matter of fact, the actual significance level of the Wald statistic is larger than that given by the chi-squared distribution (Kobayahsi, 1986). Unfortunately, the exact distribution of the test statistic is a complex function of the regressor variables and the error variances such that the exact significance level of a test score is almost impossible to obtain. However, Kobayashi (1986) shows that the distribution of W/k (that is, the Wald statistic divided by the number of regressors) is asymptotically bounded by the distribution of two F variates; $F(k, N_1 + N_2 - 2k)$ and $F(k, \min(N_1 - k, N_2 - k))$. The actual probability of observing a particular Wald statistic will lie between the bounds defined by these two variates.

Tables 7 and 8 present a series of pairwise comparisons of parameters for the clusters defined by neighbourhood socioeconomics and property attributes respectively. To be conservative, the Wald statistics are based upon contrasts in only the continuous parameters of the models (including the constant, garage and central heating dummy variables). The *p*-values presented in these tables are the upper bound of the range identified by Kobayashi. Again, these will tend to favour acceptance of the hypothesis of equality in parameters.

Nevertheless, for all comparisons in both partitions, the test statistics are significant at a greater than 95% level of confidence[38]. In accordance with theory, there are significant differences between the prices that characterise the localities on the hedonic price surface isolated in the different clusters.

[37] Parameters unique to one of the models being tested were dropped from the calculation of the statistic.

[38] Wald tests based on contrasts in all the parameters of the model are significant at a greater than 99% level of confidence for all comparisons.

Table 7: Wald test chi-squared statistics for differences between hedonic price functions for neighbourhood socioeconomic characteristics partition

	Wald Test Statistics (p-values – Kobayashi's upper bound)					
Submarket	**1**	**2**	**3**	**4**	**5**	**6**
2	118.917 (0.000)					
3	78.426 (0.000)	65.259 (0.000)				
4	74.288 (0.000)	54.36 (0.001)	55.508 (0.001)			
5	41.065 (0.024)	72.541 (0.000)	43.228 (0.014)	74.035 (0.000)		
6	50.252 (0.002)	65.859 (0.000)	86.774 (0.000)	65.997 (0.000)	48.79 (0.003)	
7	70.708 (0.000)	64.644 (0.000)	50.467 (0.003)	73.952 (0.000)	55.005 (0.001)	49.07 (0.003)

Table 8: Wald test chi-squared statistics for differences between hedonic price functions for property attribute partition

	Wald Test Statistics (p-values – Kobayashi's upper bound)				
Submarket	**1**	**2**	**3**	**4**	**5**
2	152.083 (0.000)				
3	96.017 (0.000)	70.865 (0.000)			
4	54.575 (0.001)	86.429 (0.000)	52.873 (0.001)		
5	92.327 (0.000)	95.03 (0.000)	55.162 (0.001)	60.795 (0.000)	
6	132.302 (0.0000	125.608 (0.0000	44.997 (0.010)	68.651 (0.000)	101.968 (0.000)

6. Comparison of Data Partitions

The Wald tests carried out in the previous section confirm that the HPF cannot be adequately approximated by a single linear regression. Rather partitioning the data and estimating a set of linear regressions one for each partition, reveals significant differences between the marginal prices of property attributes in different clusters. One further comparison needs to be made, that between the different partitions of the data. We wish to test which of the two partitions of the data is better at isolating those regions of the hedonic price surface between which marginal prices differ significantly.

In effect we have two competing economic theories that imply different linear regression models. For example, the set of M^a linear regressions estimated for the clusters defined by partitioning according to the attributes of properties is equivalent to the single linear regression;

$$\begin{bmatrix} \ln(P_1) \\ \ln(P_2) \\ \vdots \\ \ln(P_{M^a}) \end{bmatrix} = \begin{bmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & X_{m^a} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{M^a} \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_{M^a} \end{bmatrix} \tag{7}$$

or more succinctly;

$$y^a = X^a \beta^a + e^a \tag{8}$$

Likewise the set of M^b linear regressions estimated for the clusters defined by partitioning according to the socioeconomics of neighbourhoods could be represented by the single linear regression;

$$y^b = X^b \beta^b + e^b \tag{9}$$

Since the data for the model in (8) is partitioned differently to that in (9) it must be the case that neither model is a special case of the other.

Goodman and Dubin (1990) were the first to propose the use of the *J*-test (Davidson and Mackinnon, 1981) in order to compare the two hypotheses defined by the specifications in (8) and (9). The *J*-test requires artificially nesting the two models by including the fitted values from one specification as an explanatory variable in the other.

Consider first model a in which the data is partitioned according to the attributes of the properties themselves. Let us suppose that partitioning the data in this way generates clusters that isolate those regions of the hedonic price surface between which the marginal prices of property characteristics differ markedly. In this case, we would expect the model in Equation (8) to fit the data very well. Imagine also, that the opposite is true of model b. That is, the partitioning defined by the socioeconomics of neighbourhoods does not isolate regions of the hedonic price surface characterised by markedly different marginal prices. We could test this hypothesis by artificially nesting the two models according to;

$$\boldsymbol{y}^a = \boldsymbol{X}^a \boldsymbol{\beta}^a + \alpha^b \,\hat{\boldsymbol{y}}^{b(a)} + \boldsymbol{e}^a \tag{10}$$

where $\hat{\boldsymbol{y}}^{b(a)}$ is the $N \times 1$ vector of fitted values from the linear regression in (9) with the observations reordered to conform with the arrangement of the observations in (8).

Now if our hypothesis were correct then we would not expect $\hat{\boldsymbol{y}}^{b(a)}$, the fitted values from the socioeconomic partitioning of the data, to add significantly to the explanatory power of the model in (8). Indeed, a simple *t*-test of the single parameter α^b, can be used as test of the hypothesis. If α^b is not statistically different from zero then we can conclude that partitioning the data by the socioeconomics of neighbourhoods adds nothing to the model that is not already captured by partitioning the data according to the attributes of properties themselves.

Of course we could also test the alternative hypothesis; that partitioning the data according to the attributes of the properties themselves adds nothing to our model of the HPF that is not captured by partitioning the data according to the socioeconomic composition of neighbourhoods. To test this hypothesis we can artificially nest the two models according to;

$$y^b = X^b \beta^b + \alpha^a \hat{y}^{a(b)} + e^b \qquad (11)$$

where $\hat{y}^{a(b)}$ is the $N \times 1$ vector of fitted values from the linear regression in (8) with the observations reordered to conform with the arrangement of the observations in (9). Again an insignificant *t*-test would allow us to accept the hypothesis that little is gained through partitioning the data according to property attributes that is not already accounted for through partitioning the data according to neighbourhood socioeconomic composition.

The results of the pair of *J*-tests defined by the models in Equations (10) and (11) are recorded in Table 9.

Table 9: J-tests of alternative partitions of data

	Coefficient (s.e.)	**p-value**
H₀: *Neighbourhood Partition does not provide information beyond that already captured by Property Partition*	0.0566 (0.0086)	0.000
H₀: *Property Partition does not provide information beyond that already captured by Neighbourhood Partition*	-0.0064 (0.0038)	0.088

In this application, the *J*-test provides a clear conclusion. Including fitted values from the socioeconomics of neighbourhood partition in the model based on partitioning the data according to property characteristics significantly improves the fit of the model; α^b is significantly different from zero at over the 99.9% level of confidence. In contrast including the fitted values from the property characteristics partition into the model based on partitioning according to the socioeconomics of neighbourhoods does not significantly improve the model; α^a is not significantly different from zero at the 95% level of confidence.

Modelling the HPF by partitioning the data according to the socioeconomics of neighbourhoods statistically dominates models defined by partitioning the data according to the attributes of properties themselves.

7. Conclusion

This Chapter has examined the implications for empirical hedonic analysis arising from recent developments in the theoretical literature. We describe two sets of theoretical models; one set of models assume households choose where to live based on the characteristics of the properties themselves, the second set of models assume that households' choice of domicile is determined by the characteristics of the equilibrium sets of people that choose to inhabit the neighbourhood in which a property is located.

Using property market data from the City of Birmingham in the UK, we test the predictions of these two different models. Our empirical analysis acknowledges two characteristics of the equilibrium market predicted by both models. First, that typically the equilibrium HPF will be highly nonlinear. Second, that the equilibrium market may be characterised by clusters of properties (neighbourhoods) exhibiting similar combinations of attributes.

Using recently developed techniques of model-based clustering we identify clusters of (1) similar properties and (2) similar neighbourhoods. In contrast, with other techniques, model-based clustering provides a framework in which we are able to make statistical deductions concerning the nature and number of clusters in the data. In both cases, in accordance with the theoretical predictions, we find clear evidence of clustering. We generate two partitionings of the data one based on clustering by property attributes, the second based on clustering by neighbourhood attributes.

Our strategy for estimating the HPF derives directly from the identification of clusters. First we note that properties categorised into the same cluster lie in close proximity to each other in certain dimensions of the attribute space. By extension, these properties must also lie close to each other in these dimensions on the hedonic price surface. We hypothesise that partitioning the data generates clusters that isolate regions of the hedonic price surface between which the marginal prices of property characteristics differ markedly. Thus, rather than employing increasingly more general econometric specifications to capture the nonlinearity of the equilibrium HPF, our estimation strategy is to avoid estimating the HPF over the entire attribute space. Rather, we fit separate price functions for the properties in each cluster thereby forming local approximations to the hedonic price surface over the attribute space spanned by the properties in each cluster.

In the application described here, we find that the HPF cannot be adequately approximated by a single linear regression. Rather partitioning the data and estimating a set of linear regressions, one for each partition, reveals significant differences between the marginal prices of property attributes in different clusters. Indeed, one of the advantages of this approach when compared to estimation strategies based on nonparametric regression, is that the parameters estimated on the various covariates can be examined for interesting contrasts across clusters. In the application described here, for example, we find that the market tends to reward ethnic homogeneity within neighbourhoods.

Finally we test to see whether one of the two proposed partitions can be said to provide a better description of the data in the model than the other. Using a *J*-test we discover that partitioning the data according to the socioeconomic characteristics of neighbourhoods, statistically dominates a model in which the data has been partitioned according to the attributes of properties. It appears that differences in property prices can better be captured by looking at the differences that exist between socioeconomically differing neighbourhoods than by examining the differences that exist between different structural types of property.

CHAPTER 7. OMITTED LOCATIONAL COVARIATES IN HEDONIC ANALYSIS: A SEMIPARAMETRIC APPROACH USING SPATIAL STATISTICS

1. Introduction

The recent history of research into the hedonic analysis of property markets has witnessed a widespread recognition of the importance of spatial processes. In the theoretical literature, models have been developed in which households' choices of residential location may depend explicitly on the sets of people that choose to live in each location (e.g. Epple and Platt, 1998; Epple and Seig, 1999; Nesheim, 2002). These models predict that in equilibrium, households will sort themselves across the urban area such that the characteristics of households living in the same neighbourhood are likely to be more similar to each other than they are to the population as a whole. This equilibrium is characterised by a HPF that maintains price differentials between locations in the urban area (Nesheim, 2002).

Likewise, in the empirical literature, there has been a growing acknowledgement that the econometric methods used to estimate HPFs from property market data should explicitly concern themselves with the spatial organisation of the data (e.g. Dubin, 1988, 1992, 1998; Can, 1992; Pace and Gilley, 1997; Can and Megbolugbe, 1997; Basu and Thibodeau, 1998; Pavlov, 2000; Bell and Bockstael, 2000; Leggett and Bockstael, 2000; Gawande and Jenkins-Smith, 2001; Ihlanfeldt and Taylor, 2001, Gibbons and Machin, 2001; Gibbons, 2001).

In particular, empirical researchers are concerned with the fact that the selling prices of properties will be positively correlated over space. In addition to the price differentials generated by the sorting processes identified in the theoretical literature, there are numerous reasons why researchers might expect prices to be spatially correlated. For example, urban areas often develop piecemeal over time. Local neighbourhoods tend to be constructed at the same time and by the same developers. Consequently, properties within neighbourhoods are likely to exhibit structural similarities not only in terms of their age but also in terms of their size, layout and interior and exterior design features. Moreover, properties in the same neighbourhood also share the same physical surroundings. As such they will have comparable access to locational amenities (e.g. schools, shops, parks, transport links etc.) and exposure to disamenities (e.g. industrial sites, landfills, air pollution, noise pollution etc.). If households value proximity to (distance from) these amenities (disamenities), then the selling prices of properties will be correlated over space.

The efforts of empirical researchers to incorporate spatial considerations into their analyses have been manifold. For example, in order to attend to the theoretical prediction that property prices may vary according to the socioeconomic composition of neighbourhoods, researchers invariably include measures of neighbourhood socioeconomics in their specification of the HPF. Alternatively, HPFs can be specified such that the marginal prices of property attributes are allowed to vary according to neighbourhood characteristics (e.g. Can, 1992, Can and Megbolugbe, 1997). In a similar vein, some researchers have sought to identify sets of socioeconomically homogeneous neighbourhoods and estimate separate HPFs for properties falling into each set (e.g. Day, 2003; Goodman and Thibodeau, 2003). Using the neighbourhood clustering identified in Chapter 6, we employ this latter estimation strategy in the empirical work presented in this chapter.

Furthermore, researchers have employed ever more sophisticated data sets that provide details of many of the structural characteristics of properties and make use of geographical information systems (GIS) to construct variables that paint a comprehensive picture of each properties' access to amenities and exposure to disamenities (e.g. Lake et al., 2000). The Birmingham data set used in this study is an example of just such a data set.

Despite these advances in data collation, it seems unlikely that any data set will be sufficiently comprehensive that it captures every aspect of property construction and location that might induce correlation in prices over space.

Typically, researchers address this problem by including locational constants that crudely describe each property's location in the urban area (e.g. properties may be categorised according to postal region or perhaps administrative or political subdivisions of the urban area). Even so, there is no guarantee that these locational constants are effective proxies for variations in the unobserved covariates. In particular, it seems unlikely that the unmeasured spatial processes will operate on the exact spatial scale as the regions defined by the locational constants. Similarly, it seems implausible to expect that these spatial processes will obey the rigid boundaries imposed by the locational constants. For example, it is more likely that a property located at the edge of its allotted region will hold more in common with properties lying just over the boundary in the adjacent region, than it will with properties on the far side of its own region (Dubin, 1992). Alternatively, some researchers include as regressors polynomial expressions in the latitude and longitude of each property (e.g. Dubin, 1992; Pace and Gilley, 1997). Whilst allowing for continuous variation in prices over the urban area, this approach will only effectively capture large-scale spatial variation in prices.

The fact remains, that any empirical specification of the regressors in a HPF is unlikely to be sufficiently comprehensive to remove all spatial effects from the data. Of course, one can test this hypothesis by examining regression residuals for spatial autocorrelation. Evidence of positive spatial autocorrelation in regression residuals is an indication of spatial processes that are not captured by the specification of the HPF. As described in Section 3 of this chapter these tests require the researcher to specify *a priori* the area over which spatial autocorrelation in the regression residuals is thought to operate. However, there is no established procedure for determining this distance. Can (1992) and Bell and Bockstael (2000), for example, simply try a variety of distances and find evidence of spatially correlated residuals in all cases.

Alternatively, a more thorough appreciation of the nature of spatial dependence in regression residuals can be obtained through construction of the spatial correlogram. In the hedonic analysis of property markets, Dubin (1988, 1992, 1998) constructs spatial correlograms for residuals by taking the average correlation in residuals at progressively larger separation intervals or distance classes. In this chapter we propose a more sophisticated approach inspired by the paper of Ellner and Seifu (2002). Here we employ a test of spatial autocorrelation of regression residuals known as Moran's I statistic (Cliff and Ord, 1972). We calculate Moran's I statistic for residuals at progressively larger separation intervals. Since the distribution of I under a null hypothesis of no spatial autocorrelation is known, it is possible to establish statistically the separation interval at which correlation of the residuals is no longer a feature of the data.

Of course, having identified spatial autocorrelation in regression residuals, the researcher is faced by the troublesome task of deciding how to proceed. As described in Section 4 of this chapter, there are, in essence, three routes that may be followed. One approach is to assume that one has data on all relevant determinants of property prices and that spatial autocorrelation of the residuals is merely an artefact of a mis-specified model. Under this assumption the prognosis is that respecifying the model will solve the problem. A second approach is to assume that the true model is the model at hand but that autocorrelation among the disturbances is due to spatial dependence in the process generating the nuisance. Again the proscribed course of action is to model that nuisance process and thereby alleviate the symptoms of spatial autocorrelation of residuals. The final approach and that championed here is to accept that there are spatial features influencing property prices that are not observed by the researcher. Whilst many of these features might be the subtle nuances of location that might adequately be handled by modelling of the nuisance process, others may be substantive spatial features whose absence from the model is likely to induce missing variable bias in the parameter estimates. For example, properties located close to an abattoir are likely to exhibit considerably deflated market prices. If proximity to abattoirs is not included as a regressor in the estimated HPF then one might conclude that the model is misspecified and that the parameter estimates are unreliable.

In a non-spatial setting the presence of omitted variables presents an almost insurmountable obstacle to the researcher. However, as pointed out by Gibbons and Machin (2001) and Gibbons (2001), where the omitted variables can reasonably be expected to be features of geographical space, a course of action suggests itself. That course of action is to account for the missing covariates through the introduction of a spatial fixed effect estimated using a nonparametric kernel regression procedure. That is, for each property the influence of its particular location on its price can be estimated as the distance-weighted average of the prices of other properties in its neighbourhood. The HPF can then be estimated by linear regression using the deviations of observed prices and regressors from their expected values at each location. Gibbons and Machin (2001) call this a *Smooth Spatial Effects* (SSE) estimator.

A question that remains is over what spatial area the data should be smoothed. As Gibbons and Machin (2001) point out, this amounts to deciding upon the bandwidth for the kernel used to smooth the data. A larger bandwidth will account for spatial processes operating over a wider area, a smaller bandwidth will account for more localised phenomena. Here we propose the use of an alternative procedure suggested for use in another context by Ellner and Seifu (2002).

Construction of the spatial correlogram for the regression residuals provides a statistical indication of the area over which spatial correlation is a feature of the data. We assume that our regression model lacks covariates that operate so as to influence property prices over this spatial scale. Our choice of spatial smoothing bandwidth is motivated by the desire to remove the impacts of these missing covariates. The procedure outlined by Ellner and Seifu (2002) involves repeated estimation of the SSE model using progressively larger bandwidths. At each iteration, Moran's I statistic is calculated to assess the degree of autocorrelation in the residuals over the spatial scale identified by the correlogram. The optimal bandwidth is selected as that bandwidth at which the computed value of I matches its expectation under the hypothesis of uncorrelated residuals. Ellner and Seifu term this the *Residual Spatial Autocorrelation* (RSA) criterion.

The rest of this chapter is organised as follows. In Section 2 we introduce the data set that forms the focus of our empirical application. In Section 3 we describe Moran's I statistic as a measure of spatial autocorrelation in regression residuals. We also describe the use of Moran's I in the construction of the spatial autocorrelogram and apply this procedure to the data. In Section 4 we briefly describe models used to account for the spatial autocorrelation of residuals and introduce the smooth spatial effects estimator. In Section 5 we apply the RSA criterion of Ellner and Seifu (2002) to the data in order to choose the optimal region over which to spatially smooth. We compare the recommendations of this procedure with that of cross-validation; an alternative procedure frequently used to select bandwidths. Finally, we apply statistical tests to determine whether the parameters of the SSE differ significantly from a model that does not account for omitted spatial covariates.

2. The City of Birmingham Data Set

The case study described in this chapter is for residential house sales in the City of Birmingham. Complete data records were successfully compiled for some 10,848 residential property transactions in 1997. Some 57 observations were excluded from the final data set for various reasons leaving 10,791 observations. Descriptions of the variables used in the hedonic analysis can be found in Chapter 4.

This chapter is particularly concerned with spatial relationships within the data. For the purposes of this research we make use of two levels of spatial organisation defined for administrative purposes in Birmingham. The first of these are termed enumeration districts (EDs), and comprise the smallest area over which census data is provided by the Office of National Statistics (ONS). Birmingham is divided into 1,940 EDs, with each ED containing an average of 191 households. EDs are gathered into larger scale political units known as wards. Birmingham contains 39 wards such that each ward comprises an average of 50 EDs and 9,500 households. The organisation of these administrative spatial units are shown in Figure 1.

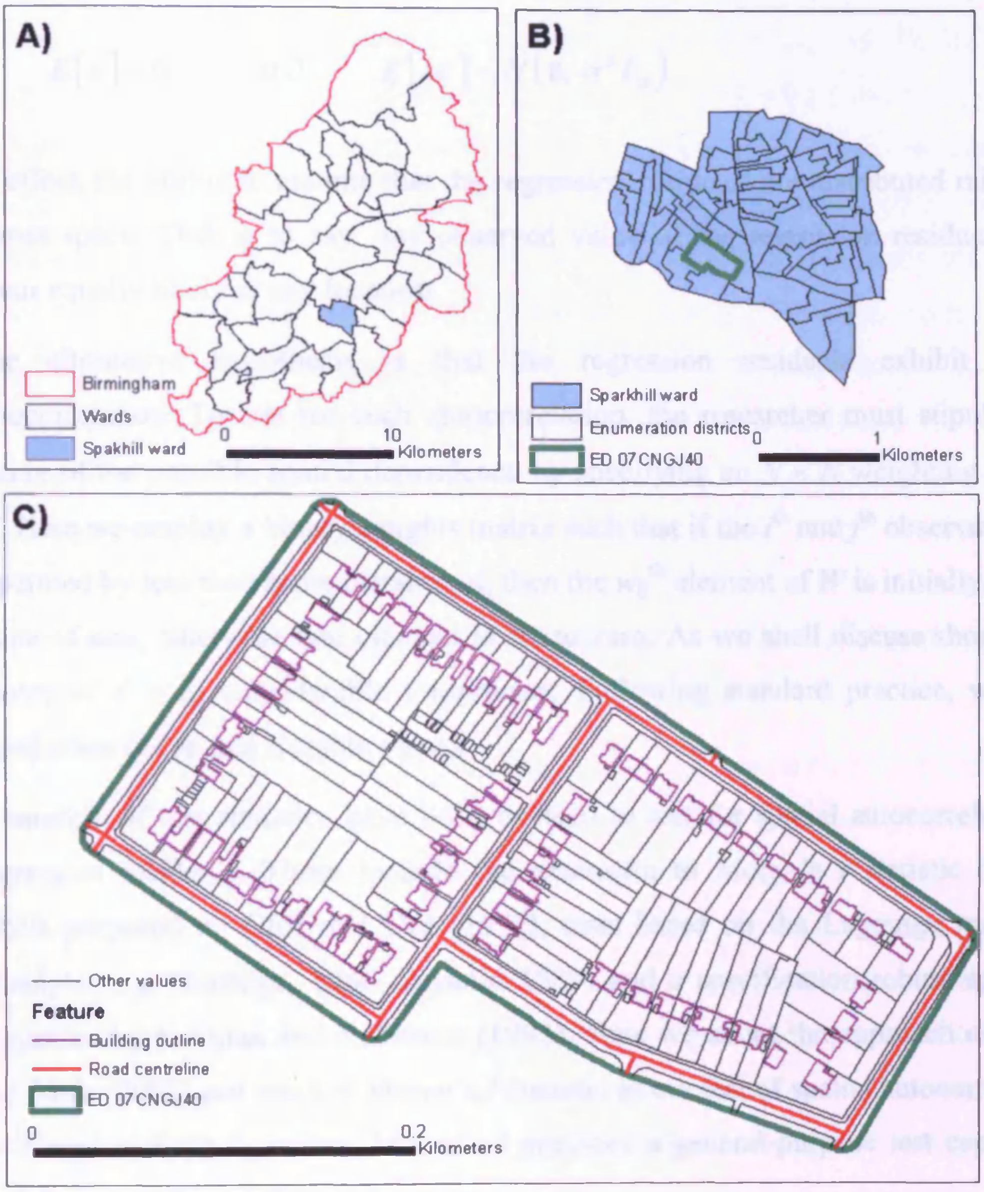
3. Assessing spatial autocorrelation in regression residuals

Following on from the analysis in the last Chapter, the Birmingham property market data is partitioned into seven clusters. Each cluster of properties is defined by the similarity of the socioeconomic composition of the neighbourhoods in which those properties are located. For each cluster we estimate the HPF as a simple linear regression;

$$\ln \boldsymbol{P}_j = \boldsymbol{X}_j \boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_j \qquad j = 1, 2, ..., M \qquad (1)$$

where j indexes clusters, $\boldsymbol{P}_j$ is the $N_j \times 1$ vector of property prices for data allocated to cluster j, $\boldsymbol{X}_j$ is the associated $N_j \times K_j$ regressor matrix, $\boldsymbol{\beta}_j$ is the $K_j \times 1$ vector of parameters and $\boldsymbol{\varepsilon}_j$ is the $N_j \times 1$ vector of regression residuals.

Figure 1: Hierarchy of administrative areas in Birmingham



Since it adds nothing to the discussion, let us simplify notation by dropping the cluster index, j. Further, to allow a more generic discussion let us replace the regressand $\ln P$ with the nonspecific vector of dependent variables $\boldsymbol{y}$, giving;

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \qquad (2)$$

Our null hypothesis is the absence of spatial autocorrelation in the regression residuals. That is we assume that;

$$E[\varepsilon] = \mathbf{0} \qquad \text{and} \qquad E[\varepsilon\varepsilon'] \sim N\left(\mathbf{0}, \sigma^2 \boldsymbol{I}_N\right) \tag{3}$$

In effect, the null is to assume that the regression residuals are distributed randomly across space. That is to say, any observed value of the regression residual could occur equally likely at any location.

The alternative hypothesis is that the regression residuals exhibit spatial autocorrelation. To test for such autocorrelation, the researcher must stipulate the nature of the possible spatial dependence by specifying an $N \times N$ weighting matrix, $\boldsymbol{W}$. Here we employ a binary weights matrix such that if the i^{th} and j^{th} observation are separated by less than some distance, d, then the $w_{ij}{}^{\text{th}}$ element of $\boldsymbol{W}$ is initially set to a value of one, otherwise that element is set to zero. As we shall discuss shortly, the choice of d is of considerable importance. Following standard practice, we row-standardise the spatial weights matrix

A number of test statistics have been devised to test for spatial autocorrelation in regression residuals. These include the extension to Moran's I statistic (Moran, 1950) proposed by Cliff and Ord (1972), tests based on the Lagrange multiplier principle (e.g. Burridge, 1980; Anselin 1988) and a specification robust approach suggested by Kelejian and Robinson (1992). Here we adopt the approach of Ellner and Siefu (2002) and employ Moran's I statistic as our test of spatial autocorrelation. As Hepple (1998) describes, Moran's I provides a general-purpose test capable of detecting most forms of spatial pattern.

Let $\hat{\varepsilon} = \ln(\boldsymbol{P}) - \boldsymbol{X}\hat{\boldsymbol{\beta}}$ be the regression residuals when $\hat{\boldsymbol{\beta}}$ is the ordinary least squares (OLS) estimator of $\boldsymbol{\beta}$, then Moran's I statistic is given by;

$$I = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} \hat{\varepsilon}_i \hat{\varepsilon}_j w_{ij}}{\sum_{i=1}^{N} \hat{\varepsilon}_i^2} \left(\frac{N}{S_0}\right) = \frac{\hat{\varepsilon}' \boldsymbol{W} \hat{\varepsilon}}{\hat{\varepsilon}'\hat{\varepsilon}} \left(\frac{N}{S_0}\right) \tag{4}$$

where N is the number of observations and $S_0 = \sum_i \sum_j w_{ij}$, the sum of all the elements in the weights matrix. The numerator in Equation (4) is a cross-products (covariance) term, while the denominator is a variance term. As such I behaves as a product-moment correlation, varying on the interval [-1,1], with 1 indicating perfect positive correlation of residuals and -1 indicating perfect negative correlation of residuals. The significance of non-zero I can be judged by comparison with the distribution of I under the null hypothesis of residuals that are randomly distributed over space.[39] Cliff and Ord (1972, 1973) showed that in large samples, this distribution was approximately normal and developed formulas for its mean and variance (see Anselin and Hudak, 1992). The statistical and analytical power of the test has been confirmed by numerous Monte Carlo studies (e.g. Bartels and Hordijk, 1977; Brandsma and Ketellapper, 1979; Anselin and Rey, 1991). Whilst Hepple (1998) has developed the exact distribution of the I statistic, here we continue to use the Cliff-Ord normal approximation due to the comparative simplicity of its calculation.

Our major concern in this section is the choice of an optimal value d to use in testing for spatial autocorrelation. That is, we wish to define a statistical procedure that indicates the area over which spatial autocorrelation of residuals is a feature of the data. As discussed in the introduction we assume that our regression model lacks covariates that operate so as to influence property prices over this spatial scale. To a greater extent, researchers in the hedonic literature have not concerned themselves with the choice of d. Indeed in testing for spatial autocorrelation, or for that matter modelling spatial autocorrelation, d is generally chosen in some *ad hoc* manner. For example, Bell and Bockstael (2000) choose a value of 600m since this is the average size of housing developments in the area. Likewise, Ihlanfeldt and Taylor (2001) choose a distance of 3 miles since this is the smallest distance to guarantee that all

[39] If we believe that the residuals are randomly distributed over space then the value of I in Equation (2) is only a single value out of a possible $N!$ values that could be found if the residuals were randomly reallocated over observations and I recalculated. Indeed, if we were to graph the density of the $N!$ possible values of I we would produce a distribution from which a standard error could be obtained. If the particular value of I is found to be a rare occurrence under randomisation then it can be inferred that some pattern of spatial autocorrelation exists in the data.

observations have at least one neighbour and because "this distance seems sufficiently large to allow for almost any type of spatial dependence".

Here we make use of the *correlogram*, more familiar to economists for its application in times series econometrics. We calculate Moran's *I* for a series of lag distances (or distance classes) from each point by specifying a weighting matrix that assigns a value of one to pairs of observations separated by a distance that falls within that class, and a value of zero otherwise. The resulting correlogram illustrates the degree of autocorrelation at each lag distance. Dubin (1988, 1992, 1998) follows a similar procedure to construct spatial correlograms for residuals from hedonic price regressions for property market data. Here however, we adopt a more sophisticated approach inspired by the paper of Ellner and Seifu (2002). We plot on the same correlogram the expected value and the 95% confidence intervals of the distribution of Moran's I under the null hypothesis of random distribution of residuals over space (using the formulas of Cliff and Ord, 1973). We take *d* as being the distance class at which the correlogram falls within the 95% confidence interval of random spatial distribution of residuals.

Our procedure differs from that of Ellner and Seifu (2002) in that they do not calculate a correlogram. Rather they plot the value of *I* for weights matrices defined by progressively larger values of *d*. We prefer our approach since the presence of substantial autocorrelation at small values of *d* may dominate the value of Moran's I statistic when calculated for more inclusive values of *d*. Thus the value of Moran's *I* statistic may remain significant at larger values of *d* even if the more distant observations bought into the calculation by extending *d* are not actually correlated.

In the application described here we estimate two specifications of the regression model in (1). In the first the regressor matrix, *X*, includes the multiplicity of structural, neighbourhood, environmental and locational variables. The correlograms for this specification are plotted in Figure 2. In the second specification we include a set of locational constants. The constants indicate in which of the 39 wards each property is located (see Figure 1). As discussed in the introduction, these wide area locational constants constitute a crude attempt to capture spatial variation in property prices that is not accounted for by the other regressors included in the hedonic analysis. The correlograms for this specification are plotted in Figure 3.

The correlograms are calculated for 100m distance classes. In Figures 2 and 3 the value of Moran's I (and its expectation and 95% confidence band under random distribution of errors) for a distance class is plotted at the upper limit of the class. The value for d, therefore, is taken as the upper boundary of the largest distance class to fall outside the 95% confidence bands such that the I statistics for successive distance classes fall consistently with these confidence bands. For example, in Figure 2 the last vertex of the correlogram for Cluster 6 to fall outside the 95% confidence bands is that for the 500m to 600m distance class. Subsequent distance classes return I statistics that are not significantly differently from what might be expected under random distribution of residuals. In this case, d is taken to be 600m. Not all cases are as clear cut. The correlogram for Cluster 5 in Figure 3 dips into the 95% confidence bands for the 300m to 400m distance class but subsequent classes return I statistics evidencing statistically significant spatial correlation. In this case d is taken to be greater than 1000m (the highest value plotted on the correlograms).

Some details of the various regressions for the two specifications and values for d are reported in Table 1 (full regression results can be found in Appendices E and F). It is immediately clear from these statistics, that including the locational constants considerably improves the specification of the model. For all seven partitions of the data the adjusted R^2 statistic is seen to increase with the inclusion of the locational constants (ranging from a minimum increase of 1.4% to a maximum of 4%, with an average across all seven clusters of 2.5%). The final two columns of Table 2 report an F-test of the significance of the locational constants. In all cases, the locational constants prove to be highly significant. These findings must be treated with caution as the F-test is only appropriate if there is no spatial autocorrelation in the residuals.

That such autocorrelation is immediately evident from the correlograms in Figures 2 and 3. For all clusters in both specifications of the model the I statistics indicate significant autocorrelation of the residuals over at least the first distance band (0 to 100m). Furthermore, comparing the correlograms in Figure 2 with those in Figure 3 underscores the importance of including wide area locational constants. In the models with no locational constants, spatial autocorrelation remains an important feature of the data even for remote distance classes. In Clusters 2, 5 and 7 for example, there is significant spatial correlation for the largest distance class plotted on the correlograms (900m to 1km).

Table 1: Summary statistics from hedonic regressions by cluster; reports *d* and an *F*-test comparing the hedonic model with and without locational constants

Cluster	N	Regressions without locational constants			Regressions with locational constants			F-test of spatial constants	
		K	Adj. R^2	d (metres)	K	Adj. R^2	d (metres)	F-stat (df)	p-value
Cluster 1	2261	64	0.685	900	96	0.709	200	6.69 (32, 2165)	<.0001
Cluster 2	1258	63	0.777	>1000	90	0.817	300	10.66 (27, 1168)	<.0001
Cluster 3	2173	63	0.776	500	96	0.791	300	5.44 (33, 2077)	<.0001
Cluster 4	895	61	0.771	200	93	0.785	100	2.64 (32, 802)	<.0001
Cluster 5	2018	63	0.751	>1000	97	0.779	200	8.27 (34, 1921)	<.0001
Cluster 6	1207	60	0.810	800	85	0.836	200	8.34 (25, 1122)	<.0001
Cluster 7	970	62	0.787	500	82	0.813	100	7.28 (20, 888)	<.0001

The introduction of the wide-area locational constants does much to improve matters. Indeed, examination of the correlograms in Figure 3 reveals that the introduction of wide-area locational constants virtually eliminates autocorrelaion for distance classes greater than 300m. However, for all clusters the correlograms reveal significant evidence of more localised autocorrelation of the regression residuals.

Figure 2: Spatial correlograms for residuals from regressions not including spatial constants

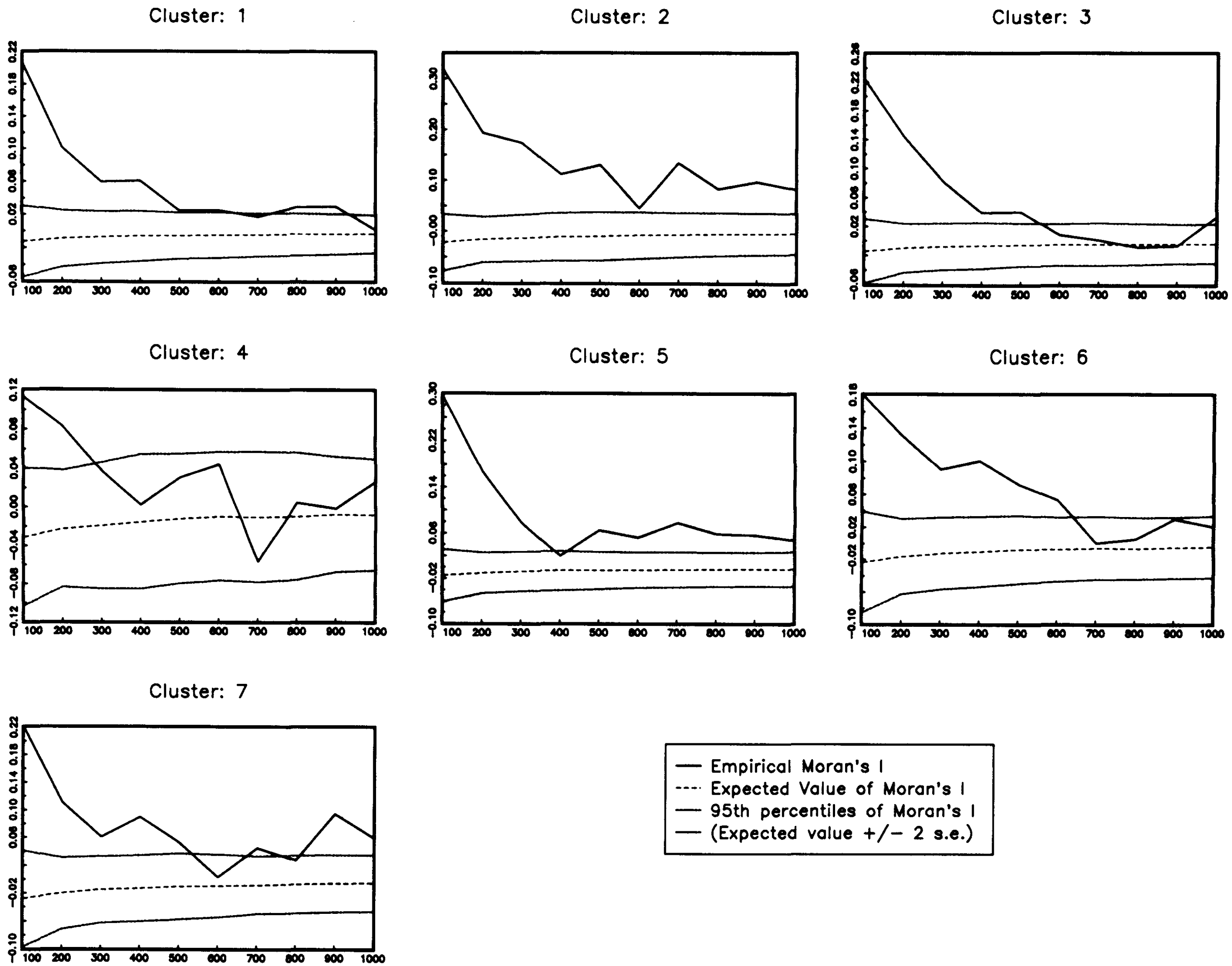
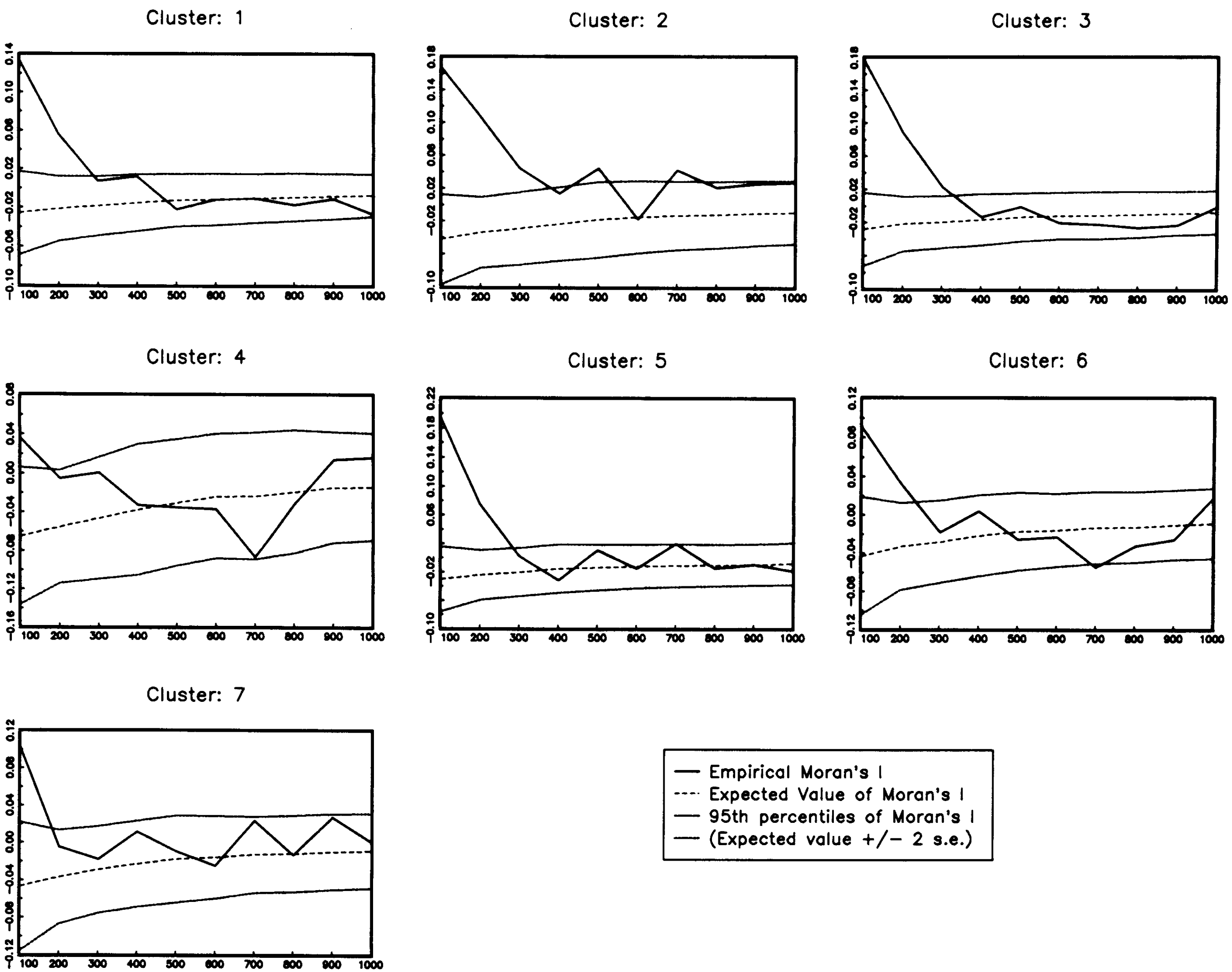


Figure 3: Spatial correlograms for residuals from regressions including spatial constants



4. Linear Regression with Spatially Correlated Residuals

The correlograms in Figure 3 reveal that the regression residuals are spatially correlated over a region of up to 300m. As such we can reject the model described by Equations (2) and (3). As described in the introduction three broad approaches to dealing with spatial autocorrelation have been proposed in the literature. We shall briefly review these in this section.

The first approach is to assume that one has data on all relevant determinants of property prices and that spatial autocorrelation of the residuals is merely an artefact of misspecifcation of the functional form of the hedonic price equation. For example, Can (1990, 1992) argues that the model in (2) is misspecified because parameter estimates are not constant over the urban landscape. Rather they are assumed to drift over space as a function of a set of regressors describing characteristics of different locations. As such, Can partitions the regressors into two sets, the $N \times K_1$ matrix $\boldsymbol{Z}_1$ and the $N \times K_2$ matrix $\boldsymbol{Z}_2$. In Can's specification $\boldsymbol{Z}_1$ comprises variables describing the socioeconomic composition of neighbourhoods whilst $\boldsymbol{Z}_2$ comprises variables describing the structural characteristics of properties. Can assumes that the parameters estimated on the $\boldsymbol{Z}_2$ regressors are not constant but vary according to the values taken by the regressors in $\boldsymbol{Z}_1$. This assumption results in what Can describes as the *spatial expansion* specification;

$$y_i = \alpha + \sum_k z_{1ki}\beta_k + \sum_l \sum_k (\gamma_{k0} + \gamma_{kl} z_{1ki}) z_{2li} + \varepsilon_i \qquad i = 1, 2, ..., N \qquad (5)$$

where i indexes property observations, y_i is i^{th} element of $\boldsymbol{y}$ (e.g. the price of the i^{th} property), z_{1ki} is the i^{th} observation of the k^{th} variable in the $\boldsymbol{Z}_1$ matrix, z_{2li} is the i^{th} observation of the l^{th} variable in the $\boldsymbol{Z}_2$ matrix and α, $\boldsymbol{\beta}$ and γ are the parameter to be estimated.

A natural extension of the spatial expansion specification is proposed by Pavlov (2000). Again, Pavlov assumes that the coefficients of a linear hedonic function vary across the urban space. However, rather than specifying a functional relationship between the spatially varying coefficients and a set of locational variables, Pavlov allows the value of the coefficients to be determined by the data. The space varying coefficients are made functions of locations according to the *space-varying coefficients* (*SVC*) specification;

$$y_i = \alpha_i(c_{1i}, c_{2i}) + \sum_k \beta_{ki}(c_{1i}, c_{2i})x_{ki} + \varepsilon_i \qquad i = 1, 2, ..., N \qquad (6)$$

where α_i is a space varying constant specific to the location of the i^{th} observation as defined by its coordinates $c_i = (c_{1i}, c_{2i})$. Likewise, β_{ki} is a space-varying coefficient specific to the location of the i^{th} observation. An estimate of the coefficients at any particular location is made using weighted least squares. Only the m nearest observations to the location of interest receive non-zero weights in this regression. Greater weight is attributed to observations more proximal to the location of interest according to the Epanechnikov weighting scheme (Epanechnikov, 1969). The *SVC* method allows for both the intercept and the slope parameters of the HPF to differ by location. However, this ability to handle spatial processes in the data comes at a cost. As Pavlov (2000) points out, the space-varying coefficients method lacks a theoretical inferential framework. Since the parameters of the hedonic vary continuously over space it is not possible to judge the statistical significance of any particular regressor in determining property prices.

Another respecification of the hedonic model that specifically accounts for spatial processes is the *spatial autoregressive* model. This model has been studied variously by Anselin (1989), Can (1992), Can and Megbolugbe (1997) and Gawande and Jenkins-Smith (2001). In the spatial autoregressive model the price of a property is deemed to be determined, in part, by the prices of neighbouring properties according to;

$$\boldsymbol{y} = \rho \boldsymbol{W}\boldsymbol{y} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \qquad (7)$$

where $\boldsymbol{X}$, $\boldsymbol{\beta}$ and $\boldsymbol{\varepsilon}$ are defined as previously, $\boldsymbol{W}$ is the $N \times N$ spatial weighting matrix, $\boldsymbol{y}$ is the $N \times 1$ vector of property prices, and ρ is the spatial autoregressive coefficient. Can (1990) argues that this specification has some merits since it mimics the actual workings of the property market in which estate agents appraise the value of a property according to both its own attributes and the price history of houses in the neighbourhood.

A second approach is to assume that the true model is the model at hand but that autocorrelation among the disturbances is due to spatial dependence in the process generating the nuisances. We call approaches that make this assumption *spatial error dependence* (SED) models. The SED approach has become increasingly popular in applied work, chiefly because of advances in the ease with which models of this type can be estimated (Kelejian and Prucha, 1999; Bell and Bockstael, 2000).

The consequence of a spatially dependent nuisance process is that the observations contain less information than if they had been independent. Indeed the statistical properties that are attributed to an estimator such as LS when errors are i.i.d. do not hold in this case. Nonetheless, the parameter estimates from the application of LS will not be biased, merely inefficient. In this case, the proscribed course of action is to model the nuisance process so as to obtain approximately the same quantity of information as provided by an independent set of observations.

For example, we might assume that the autocorrelation follows the first order Markovian scheme;

$$\boldsymbol{\varepsilon} = \lambda \boldsymbol{W} \boldsymbol{\varepsilon} + \boldsymbol{u} \tag{8}$$

or equivalently;

$$\boldsymbol{\varepsilon} = \left(\boldsymbol{I}_N - \lambda \boldsymbol{W}\right)^{-1} \boldsymbol{u} \tag{9}$$

where λ is the error dependence parameter and $\boldsymbol{u}$ is the usual $N \times 1$ vector of random error terms with expected value zero and variance-covariance matrix $\sigma^2\boldsymbol{I}$. Notice that $\lambda = 0$ implies $\boldsymbol{\varepsilon} = \boldsymbol{u}$ and there is no spatial dependence in the data. This particular model has been studied by various authors including Pace and Gilley (1997), Kelejian and Prucha (1999), Bell and Bockstael (2000) and Leggett and Bockstael (2000). Along similar lines, Dubin (1988, 1992, 1998) and Basu and Thibodeau (1998) develop explicit models of the nuisance process and simultaneously estimate the parameters of this process along with the regression coefficients using maximum likelihood.

Of course, SED models impose considerable structure on the processes determining spatial correlation in regression residuals. For example, they assume isotrophy. That is they assume the same model of error dependence can be applied over all space. Furthermore, spatial autocorrelation of the regression residuals is induced by locational features influencing property prices that are not observed by the researcher. SED models assume that these comprise the subtle nuances of location that might adequately be handled by modelling the nuisance process. Alternatively, the omitted spatial covariates may be substantive features whose absence from the model is likely to induce missing variable bias in the parameter estimates.

In a non-spatial setting the presence of omitted variables presents an almost insurmountable obstacle to the researcher. However, as pointed out by Gibbons and Machin (2001) and Gibbons (2001), where the omitted variables can reasonably be expected to be features of geographical space, a course of action suggests itself. Gibbons and Machin (2001) propose the *smooth spatial effects* (SSE) estimator which they specify as;

$$y_i = \boldsymbol{x}_i\boldsymbol{\beta} + q(\boldsymbol{c}_i) + \varepsilon_i \qquad i = 1, 2, \ldots, N \tag{10}$$

where $\boldsymbol{x}_i$ is the vector of observed regressors for the i^{th} observation, $\boldsymbol{c}_i = (c_{1i}, c_{2i})$ is the coordinates vector establishing the location of the i^{th} observation in space and $q(\cdot)$ is some unknown function. In effect, the specification in (10) replaces the unobserved spatial covariates with an element that is a function of location. The influence of these unobserved covariates on property prices is determined by the unknown function $q(\cdot)$. Since the influence of $q(\cdot)$ on property prices is handled nonparametrically the SSE presents an extremely flexible approach to dealing with unobserved spatial covariates.

Equation (9) is a specific example of a more general class of semiparametric models known as *partially linear models*. In that context, Robinson (1988) shows that (10) can be rewritten as;

$$y_i - E[y \mid \boldsymbol{c}_i] = (\boldsymbol{x}_i - E[\boldsymbol{x} \mid \boldsymbol{c}_i])\boldsymbol{\beta} + \varepsilon_i \tag{11}$$

suggesting that $\boldsymbol{\beta}$ can be estimated in a two-step procedure;

- First, the unknown conditional means $E[y \mid \boldsymbol{c}_i]$ and $E[\boldsymbol{x} \mid \boldsymbol{c}_i]$ are estimated using a nonparametric estimation technique.
- Second, the estimates are substituted in place of the unknown functions in Equation (10) and ordinary regression techniques employed to estimate $\boldsymbol{\beta}$.

Indeed, Robinson shows that the resulting parameter estimates are asymptotically equivalent to those that would be derived if the true functional form of $q(\cdot)$ were known and could be used in the estimation. That is, estimating Equation (10) is asymptotically equivalent to knowing both the values taken by the missing spatial covariates and knowing how the way in which these covariates impact on property prices.

In the hedonic literature, Robinson's model has been employed in a slightly different context by Anglin and Gençay (1996). Both Gibbons and Machin (2001) and Anglin and Gencay (1996) employ the Nadaraya-Watson nonparametric estimator to determine the quantities $E[y \mid \boldsymbol{c}_i]$ and $E[\boldsymbol{x} \mid \boldsymbol{c}_i]$. Notice that these quantities are simply the expected values of y and $\boldsymbol{x}$ at a particular location. In effect, the Nadaraya-Watson estimator calculates these expectations by taking the weighted average of the values of observations close to that location. Whether an observation is considered close to the location is determined by the bandwidth parameter b. The larger the value taken by b, the more observations are drawn into the calculation of the average. Further, the weight allotted to each observation in the calculation of the local average is determined by the kernel function. The kernel function must be symmetric, continuously differentiable and integrate to unity. Moreover, most commonly used kernel functions allot greater weight to observations that are in close proximity to the location than to those that are further away.

An alternative is to employ *local linear* estimators which offer significant gains over the Nadaraya-Watson estimator especially at the boundaries of the data and when the data is not equally spaced (see Fan, 1992 or Hastie and Loader, 1993, for more detailed discussion). Furthermore, as we shall discuss shortly, our estimation strategy requires repeated nonparametric estimation of the quantities $E[y \mid \boldsymbol{c}_i]$ and $E[\boldsymbol{x} \mid \boldsymbol{c}_i]$. Since nonparametric regression can be extremely time-consuming and computer-intensive, we employ fast implementation techniques as described in Fan and Marron (1994), Wand (1994) and Bowman and Azzalini (2003).

In particular, we begin by summarising the density of observations over space by linearly binning onto the verteces of a regular spatial grid. In this application the margins of the cells of the grid are set to 150m. Likewise we summarize the values of y and each of the variables present in $\boldsymbol{x}$ by calculating their linearly weighted averages at each of the verteces of the grid. Furthermore, to take advantage of computational savings offered by the use of the fast Fourier transform (FFT) we choose to use a bivariate Gaussian kernel function. Given a choice of smoothing bandwidth b, the expected values of y and $\boldsymbol{x}$ are calculated at each vertex of the grid using local linear regression. Finally, the values at each particular property location are recovered by linearly interpolating from the values of the four most proximate verteces of the grid. Since the data set is relatively large, binning the data and employing FFT-based calculations was found to be many times quicker than employing a naïve implementation of local linear regression.

5. Choice of Spatial Smoothing Parameter

A question that remains is the choice of smoothing bandwidth, b. A larger bandwidth will account for spatial processes operating over a wider area, a smaller bandwidth will account for more localised phenomena. For example, observe Figures 4 and 5. These provide plots of $E[y \mid \boldsymbol{c}_i]$, that is the expected value of $\ln P$ at a given location, for two different smoothing bandwidths. Notice first that in both cases, there is considerable variation in $E[y \mid \boldsymbol{c}_i]$ across the urban area. In particular, notice the substantial peak in the north-east section of the plots (that is, towards the back right of the cube in the figures). These peaks correspond to the desirable north-eastern suburbs of the City of Birmingham. Notice further that using a larger bandwidth as in Figure 4, results in a simpler, less convoluted surface than using a smaller bandwidth as in Figure 5. The smaller bandwidth, brings to light possibly important local features of the data that may be masked by the use of a larger bandwidth.

Gibbons and Machin (2001) choose a bandwidth motivated by the concern that spatially smoothing the data over too small an area will impact upon the parameter estimate for the variable that forms the focus of their study (namely, proximity to primary schools). Here we select a bandwidth using the *Residual Spatial Autocorrealtion* (RSA) criterion suggested for use in a slightly different context by Ellner and Seifu (2002).

Figure 4: Spatial smoothing with 1800m bandwidth (plotted on a 600m grid from the South-West)

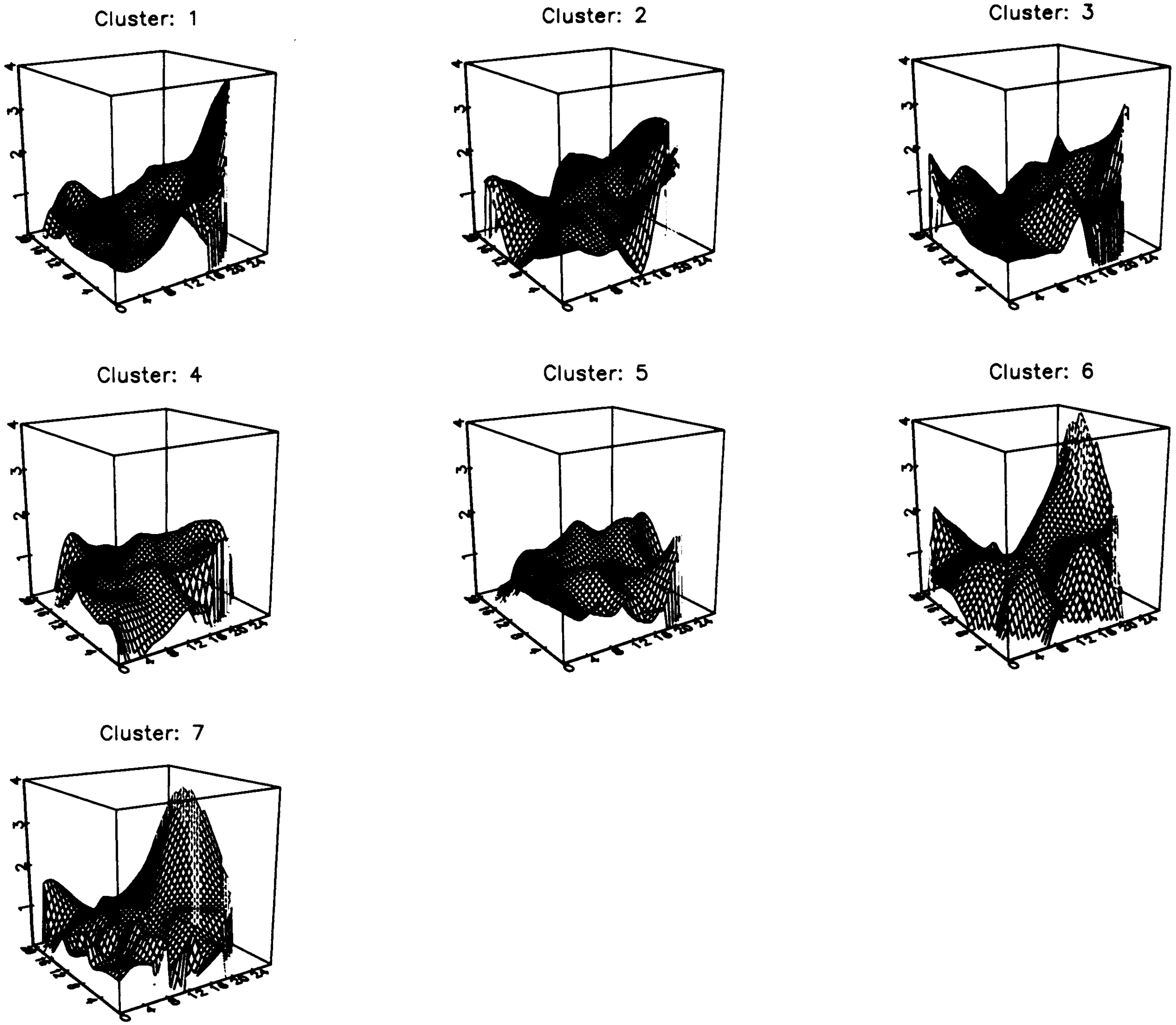
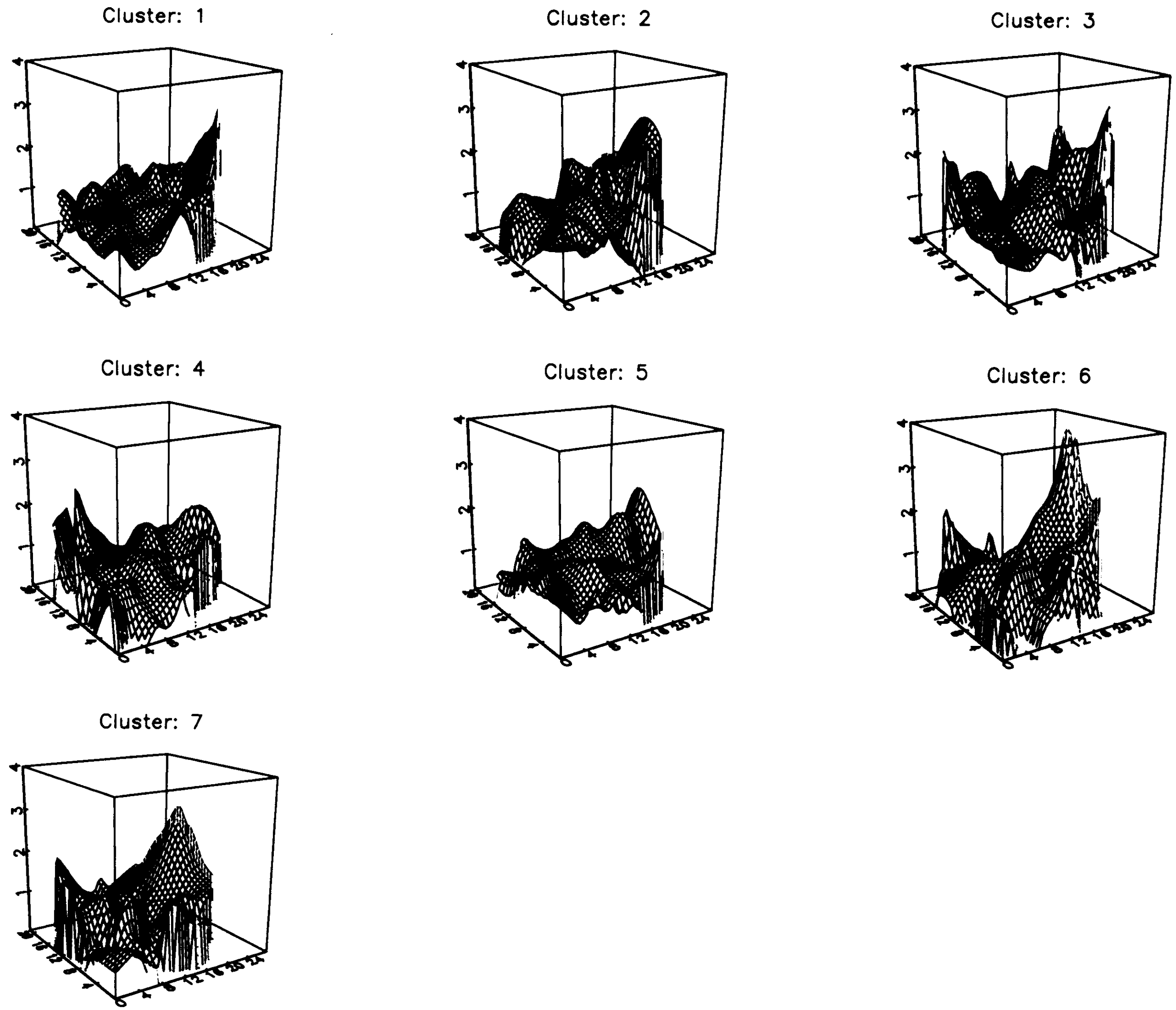


Figure 5: Spatial smoothing with 1200m bandwidth (plotted on a 600m grid viewed from the South-West)



The logic behind Ellner and Seifu's procedure is simple. In Section 3 we discussed how spatial statistics could be used to assess optimal d; that is, the area over which spatial autocorrelation of residuals is a feature of the data. Having established that the residuals show evidence of spatial autocorrelation, we conclude that our regression model lacks covariates that operate so as to influence property prices over the spatial scale given by d. Consequently, the Ellner and Seifu procedure is to search across different smoothing bandwidths, b, and for each bandwidth calculate Moran's I statistic to test for spatial autocorrelation of the residuals over an area d. Acceptable smoothing bandwidths are those for which we can reject the hypothesis of spatial autocorrelation in the residuals.

Figure 6 plots the value of Moran's I statistic for values of b at 25m intervals between 300m and 1,800m for each cluster. Also plotted in Figure 6 are the expected values of Moran's I and the 95% confidence intervals for the statistic under the assumption of randomly distributed residuals. The optimal spatial smoothing bandwidth is chosen as that at which Moran's I statistic is approximately equal to its expected value. This bandwidth is reported in Table 2 along with the upper and lower values for b at which it is still possible to reject the hypothesis of spatially autocorrelated residuals.

A commonly applied alternative for choosing bandwidths is cross-validation a selection criterion which has been shown to have some asymptotic optimality features (Härdle and Marron, 1985). For example, cross-validation was applied by Anglin and Gencay (1998) in choosing the degree of smoothing in their semiparametric estimator for a hedonic price model. Details of the cross-validation procedure for selecting bandwidth can be found in Chapter 5.

Accordingly, we calculate the cross-validation statistics for various bandwidths and plot these alongside the I statistics from the RSA procedure in Figure 6. The optimal bandwidth according to the cross-validation procedure is that which minimises the cross-validation statistic. These values are recorded in Table 2.

Notice that in four of the clusters (2, 3, 4 and 6), the bandwidth selected using cross-validation falls within the 95% confidence bounds of Moran's I. That is, if we were to smooth the data in these clusters using the optimal cross-validation bandwidth we would find that we could reject the hypothesis of autocorrelation in the residuals over an area of radius d. In the remaining three clusters (1, 5 and 7) cross-validation indicates a higher value for b than would be choosen by selecting a bandwidth that eradicates spatial autocorrelation in the residuals over an area of radius d.

Table 2: Comparison of bandwidth choice using RSA and cross-validation criteria

Cluster	Bandwidth (metres)				Hausman Test		
	RSA (lower)	RSA (upper)	RSA (optimal)	Cross-Validation	Statistic	df	p-value
Cluster 1	425	725	550	1,050	65.72	52	0.067
Cluster 2	525	875	675	525	46.74	48	0.525
Cluster 3	450	750	600	575	60.03	52	0.208
Cluster 4	375	1800	650	1,150	31.09	46	0.955
Cluster 5	300	550	450	775	75.93	50	0.010
Cluster 6	450	950	625	600	41.41	48	0.738
Cluster 7	<300	850	400	1,025	29.94	47	0.975

We test to see whether choosing the bandwidth through cross-validation rather than by the RSA criterion makes a difference. As pointed out by Gibbons (2001) sensitivity to bandwidth choice can be tested by the usual Hausman test for equivalence of parameters in alternative estimators. Denote by $\hat{\beta}^w$ the estimator using the wider bandwidth that is consistent under both the null and the alternative hypotheses, and by $\hat{\beta}^n$ the estimator using the narrower bandwidth that is fully efficient under the null but inconsistent if the null is not true. The Hausman statistic is given by;

$$\tau_H = \left(\hat{\beta}^n - \hat{\beta}^w\right) Var\left(\hat{\beta}^n - \hat{\beta}^w\right)^{-1} \left(\hat{\beta}^n - \hat{\beta}^w\right) \tag{12}$$

As Hausman (1978) shows, under the null hypothesis, the middle term in (12) (the variance matrix of the vector of differences between the parameters of the two estimators) asymptotically reduces to $Asy.Var\left(\hat{\beta}^n\right) - Asy.Var\left(\hat{\beta}^w\right)$. Of course, to make use of the Hausman result one must be able to consistently estimate the asymptotic variance matrices of the two sets of parameter estimates under the null. Unfortunately, in the presence of spatial autocorrelation of unknown form such an estimate is unavailable. Consequently, we apply a bootstrap procedure to estimate $Var\left(\hat{\beta}^n - \hat{\beta}^w\right)$. We sample with replacement from the unsmoothed data and re-estimate the SSE model using the bandwidths implied by first the RSA criterion and then cross-validation. For each bootstrap sample we calculate the difference between the two vectors of parameters. The desired variance matrix is estimated by calculating the empirical variance matrix of the differences resulting from 1,000 replications of the bootstrap procedure.[40]

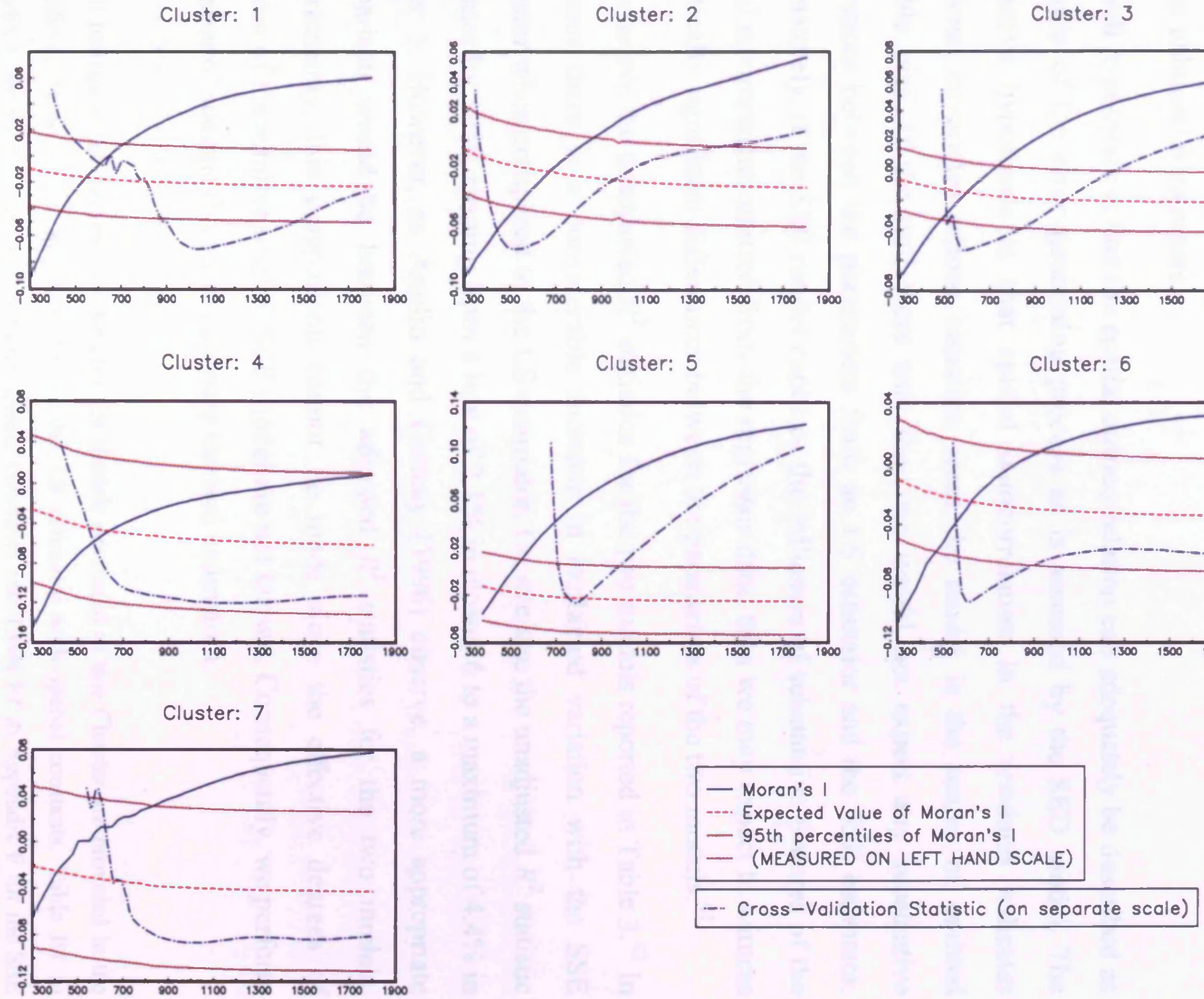
The Hausman test statistics reported in Table 2 are based on a subset of the regression parameters. We do not include the parameters for locational constants in the tests since these prove to be somewhat unstable in the SSE model where much of their influence is obviated by spatial smoothing of the data. Furthermore, the model specification includes numerous sets of dummy variables detailing categorical descriptors of property attributes (e.g. numbers of bathrooms, bedrooms, storeys etc.). When particular categories in these dummy variable sets are poorly represented in the data, it may prove impossible to estimate all the parameters of the model for every bootstrap sample. The tests are based on all parameters that are successfully estimated for each iteration of the bootstrap.

The Hausman test reveals that significant differences in the parameters can be discerned in only two of the clusters; Cluster 1 with over 10% confidence and Cluster 5 with over 5% confidence. In general then, our data suggests that choosing a bandwidth using the RSA criterion does not result in parameter estimates that differ significantly from those estimated using a bandwidth selected using cross-validation. Nonetheless, we contend that the RSA criterion provides an intuitive criterion by which bandwidths can be selected and, through the elimination of spatial autocorrelation, permits statistical inference and testing to proceed using standard econometric tools whilst imposing little assumed structure on the model of the HPF.

---

[40] Since we are estimating a variance matrix and not the tails of a distribution (as is usually the case with bootstrap procedures) we do not require a very large number of replications. Even with 1,000 replications the bootstrap took nearly 12 hours to run for each cluster on a PC with a 2.8 GHz Intel Pentium 4 processor with 512 Mb of RAM. The bootstrap would have been unfeasible without the application of the fast local linear regression procedures described in Section 4.

Figure 6: Moran's *I* and cross-validation statistics at various spatial smoothing bandwidths by cluster



6. Evidence of Omitted Locational Covariates

The analysis of the previous sections has determined that the regression residuals from LS estimation exhibit patterns of spatial autocorrelation. The final comparison we wish to make is between the SSE estimator (selecting bandwidth using the RSA criterion) and an SED estimator. Under the assumptions of the SED model, LS returns unbiased parameters

Our null hypothesis is that the spatial autocorrelation can adequately be described as a feature of the error generating process as is assumed by the SED model. The alternative hypothesis is that spatial autocorrelation in the residuals indicates locational covariates whose omission from the model is the source of omitted variable bias. If the null were true then we would not expect any substantive differences between the parameters from an LS estimator and the SSE estimator. Alternatively, if the SSE model captures the influence of substantive features of the spatial environment omitted from the regressor data, then we may expect to witness statistically significant differences between the parameters of the two models.[41]

First observe the unadjusted R^2 statistics for the two models reported in Table 3.[42] In all cases there is a considerable increase in explained variation with the SSE estimator when compared to the LS estimator. On average the unadjusted R^2 statistic increases by 3.1%, ranging from a low of 2.1% in cluster 6 to a maximum of 4.4% in cluster 5. However, as Anglin and Gencay (1996) observe, a more appropriate comparison would be between the adjusted R^2 statistics for the two models. Unfortunately, this comparison cannot be made since the effective degrees of freedom of the semiparametric SSE model are not known. Consequently, we perform a number of statistical tests to compare the two estimators.

---

[41] Full listings of the parameter estimates for models presented in this Chapter are provided in the Appendices; Table D1 in Appendix D for the LS estimator with spatial constants, Table E1 in Appendix E for the LS estimator without spatial constants and Table F1 in Appendix F for the SSE estimator.

[42] Following Anglin and Gençay (1996) we calculate the R^2 statistic as $R^2 = \hat{y}'\hat{y}/y'y$ where each element of $\hat{y}$ is given by $\hat{y}_i = \hat{E}[y \mid c_i] + \left(x_i - \hat{E}[x \mid c_i]\right)\hat{\beta}$ (where a circumflex denotes an estimated quantity).

Table 3: Unadjusted R^2 statistics from LS and SSE models

Cluster	R^2 OLS	R^2 SSE
Cluster 1	0.721	0.760
Cluster 2	0.830	0.864
Cluster 3	0.800	0.827
Cluster 4	0.807	0.838
Cluster 5	0.790	0.834
Cluster 6	0.847	0.868
Cluster 7	0.829	0.853

Our testing strategy is to compare the SSE model to the LS model under the null hypothesis that the LS model is correctly specified though there may remain spatial autocorrelation in the nuisance process.

Following, Robinson (1988) and Anglin and Gencay (1996) we first apply the Hausman test.

$$\tau_H = \left(\hat{\beta}^{SSE} - \hat{\beta}^{LS}\right) Var\left(\hat{\beta}^{SSE} - \hat{\beta}^{LS}\right)^{-1} \left(\hat{\beta}^{SSE} - \hat{\beta}^{LS}\right) \tag{13}$$

Once again, we are unable to use Hausman's expression for the variance of the difference in the two parameter vectors since we do not have an easy way of computing the asymptotic variance matrix of the LS estimator when the residuals are spatially correlated. Instead we employ a bootstrap procedure identical to that outlined in Section 5. The test statistic, τ_H is asymptotically distributed chi-squared with degrees of freedom equal to the number of parameters estimated by both models.

A second test is that proposed by Whang and Andrews (1993). Their test is based on the vector of sample moments;

$$\boldsymbol{r} = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - E[y \mid \boldsymbol{c}_i] - (\boldsymbol{x}_i - E[\boldsymbol{x} \mid \boldsymbol{c}_i])'\hat{\beta}\right)(\boldsymbol{x}_i - E[\boldsymbol{x} \mid \boldsymbol{c}_i]) \tag{14}$$

Clearly, if $\hat{\beta}$ in (11) is replaced by $\hat{\beta}^{SSE}$ then $\boldsymbol{r}$ will be a vector of zeros since (14) is simply the set of normal equations for the SSE estimator. Of course, under the null $\hat{\beta}^{LS}$ should be approximately equal to $\hat{\beta}^{SSE}$. As such the Whang and Andrews test requires that $\hat{\beta}$ in (14) be replaced by $\hat{\beta}^{LS}$. If the null holds then the moments in (14) should still approximate a vector of zeros. The test statistic is given by;

$$\tau_{WA} = \boldsymbol{r}'\hat{\Phi}^{-1}\boldsymbol{r} \tag{15}$$

where $\hat{\Phi}$ is a consistent estimator of the variance matrix of $\boldsymbol{r}$ under the null. Whang and Andrews (1993) show that τ_{WA} is asymptotically distributed chi-squared with degrees of freedom equal to the number of parameters estimated by the SSE model. Furthermore, Whang and Andrews (1993) give formulas for $\hat{\Phi}$ when the residuals are correlated. Here we prefer to bootstrap $\hat{\Phi}$ by resampling with replacement from the original data 1,000 times, re-estimating the LS and SSE models and calculating $\boldsymbol{r}$ for each bootstrap sample. Our bootstrap estimate of $\hat{\Phi}$ is the empirical variance matrix of the 1,000 bootstrap estimates of $\boldsymbol{r}$.

As discussed in Section 5, the Hausman and Whang and Andrews test statistics are based on a subset of the regression parameters. Again we do not include the parameters for locational constants in the tests, nor can we include parameters that are not estimated for every bootstrap sample.

The final test applied here is that of Li and Wang (1998). Similar to the Whang and Andrews test, the Li and Wang test statistic is based upon the residual from a "mixed" regression;

$$u_i = y_i - \hat{\beta}_0^{LS} - \boldsymbol{x}_i\hat{\boldsymbol{\beta}}^{SSE} \tag{16}$$

where $\hat{\beta}_0^{LS}$ is the estimated constant from the LS regression. Their test statistic is based upon a standardised kernel estimator of the moment condition $E[u_i E[u_i \mid \boldsymbol{c}_i] f(\boldsymbol{c}_i)]$, where $f(\boldsymbol{c}_i)$ is the spatial density of the observations.

The test statistic is given by;

$$\tau_{LW} = \frac{\sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N} u_i\, u_j\, K_b\left(c_i - c_j\right)}{\left(\sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N} 2u_i^2\, u_j^2\, K_b^2\left(c_i - c_j\right)\right)^{1/2}} \tag{17}$$

where $K_b\left(c_i - c_j\right)$ is a kernel function. Li and Wang (1998) show that under the null the test statistic is distributed $N(0,1)$.

The results of these tests are presented in Table 4.

Table 4: Test comparing LS null against SSE alternative

Cluster	Hausman		Whang & Andrews		Li & Wang	
	Stat (df)	***p*-value**	**Stat (df)**	***p*-value**	**Stat**	***p*-value**
Cluster 1	85.47 (52)	0.002	112.74 (51)	0.000	4.984	0.000
Cluster 2	91.77 (48)	0.000	136.80 (48)	0.000	4.177	0.000
Cluster 3	91.32 (52)	0.001	95.81 (52)	0.000	1.476	0.070
Cluster 4	34.37 (46)	0.896	51.00 (44)	0.284	-1.463	0.928
Cluster 5	98.41 (50)	0.000	143.62 (50)	0.000	5.505	0.000
Cluster 6	39.82 (48)	0.794	56.78 (48)	0.180	-1.418	0.922
Cluster 7	43.42 (47)	0.621	58.03 (47)	0.130	0.252	0.401

All three tests support the same conclusion. In clusters 4, 6 and 7 we cannot reject the hypothesis that the LS model is correctly specified when compared to an SSE alternative. In these cases, one is safe to assume that the SED model provides an effective model of the spatial processes in operation. This result is supported by the evidence of the correlograms in Figure 2 which showed that spatial autocorrelation was weakest in these three clusters.

In contrast, for clusters 1, 2, 3 and 5 we can unequivocally reject the null. In these cases all three tests support the conclusion that the LS model is incorrectly specified and that the SSE model returns significantly different parameter estimates. In short, these clusters show strong evidence of the presence of omitted spatial covariates. Applying a SED estimator to these data would provide biased estimates of the model parameters.

7. Conclusions

Spatial autocorrelation of regression residuals is a common feature of many econometric models. For example, here we find strong evidence for spatial autocorrelation in the regression residuals from a hedonic model that examines differences in property prices in the City of Birmingham in the United Kingdom.

To gain a more thorough appreciation of the nature of spatial dependence in regression residuals we propose construction of a spatial correlogram plotting Moran's I statistic for residuals at progressively larger separation intervals. Since the distribution of I under a null hypothesis of no spatial autocorrelation is known, it is possible to establish statistically the separation interval at which correlation of the residuals is no longer a feature of the data. In this case, we find that this separation interval differs between various subsets of the data, ranging from 100m to 300m.

Over recent years, a substantial literature has arisen concerning itself with how best to estimate econometric models blighted by spatial autocorrelation. In general, the preferred approach has been to assume that spatial correlation in regression residuals is the consequence of some modelable process generating the nuisances. Maximum likelihood and general method of moments estimators have been proposed for such SED models.

Of course, spatial autocorrelation of the regression residuals is induced by spatial features influencing property prices that are not observed by the researcher. The SED models assume that spatial autocorrelation is a consequence of an amalgam of the many subtle nuances of location and that this amalgam might adequately be regarded as a nuisance process.

However, the possibility exists that the researcher fails to observe substantive spatial features whose absence from the model is likely to induce missing variable bias in the parameter estimates. Fortunately, where the omitted variables are expected to be features of geographical space, a course of action suggests itself. In particular, we employ the SSE estimator of Gibbons and Machin (2003) and Gibbons (2003). The SSE accounts for missing spatial covariates by nonparametrically smoothing the data over a proscribed area.

In our application we spatially smooth the data using local linear regression. This approach offers significant gains over the Nadaraya-Watson smoother, especially at the boundaries of the data and when the data is not equally spaced. Furthermore, we adopt Ellner and Seifu's (2002) RSA criterion in order to select the spatial smoothing bandwidth. The RSA criterion selects that spatial bandwidth which eliminates spatial autocorrelation from the regression residuals. We compare this to the bandwidth selected through the minimisation of the cross-validation statistic, a selection criterion which has been shown to have some asymptotic optimality features (Härdle and Marron, 1985). In most cases we find that we cannot reject the hypothesis that the parameters from the SSE model using bandwidths suggested by the RSA criterion and cross-validation are equal. We contend that the spatial smoothing bandwidth for the SSE model should be selected using the RSA procedure. In particular, this procedure provides an intuitive criterion by which bandwidths can be selected and through the elimination of spatial autocorrelation, permits statistical inference and testing to proceed using standard econometric tools whilst imposing little assumed structure on the model of the HPF.

Finally, we have applied statistical tests to determine whether the parameters of the SSE estimator differ significantly from those of a SED estimator. In cases where the correlogram of the regression residuals indicates that spatial autocorrelation is an important feature of the data, we find that we can clearly reject the hypothesis that the two estimators return the same parameter estimates. In these cases, we have strong evidence for the presence of substantive omitted spatial covariates, such that the application of an SED estimator would provide biased estimates of the model parameters.

PART 3

EMPIRICAL ESTIMATION OF HOUSEHOLD PREFERENCES

CHAPTER 8. THE THEORY OF WELFARE ANALYSIS FROM HEDONIC MARKET DATA

1. Introduction

As described in Chapter 2, our interest in hedonic markets stems from the fact that environmental quality can be counted amongst the attributes of a property and as such will be reflected in a property's price. When households make decisions about where to live, they are also making a decision that forces them to trade off between money and environmental quality. The final part of this thesis concerns itself with using information on these decisions to identify household preferences for environmental quality.

In Chapter 2, we described how one particular description of a household's preferences, the bid function, was especially useful. In particular, following Bartik's (1988) analysis, the bid function was shown to provide the information needed to calculate a monetary measure of the change in economic welfare resulting from a change in environmental quality.

In this Chapter, we focus on the issue of deriving estimates of the bid function from real world property market data. As shall become evident, this is not as simple a task as might be hoped.

2. The Marginal Bid Function

The bid function, $\theta(z; y, s, u)$ describes the amount of money that a household would be prepared to pay for a property with attributes z in order to enjoy the level of utility, u. Of course, the amount that a household would bid for a particular property will not depend solely on the level of utility specified in the bid function. Rather, the household's income, y, and socioeconomic characteristics, s, will also influence their bid.

As shown in Chapter 1, the bid function can be illustrated as bid curves. Bid curves depict combinations of property attributes, z, and payments for those attributes, θ, between which the household is indifferent (i.e. combinations that confer the same utility on the household).

For our present purposes, it frequently proves more convenient to work with the marginal bid function. That is, a function that shows how much a household is willing to pay for each extra unit of housing attribute z_i, so as to maintain the same level of utility, u. Mathematically the marginal bid function is the partial derivative of the bid function. Remember from Equation (15) of Chapter 1 that the bid function is defined as;

$$\theta(z;\ y,s,u) = y - x(z;\ s,u) \tag{1}$$

Thus the marginal bid function is given by;

$$b_{z_i}(z_i; z_{-i}, s, u) = \frac{\partial\theta(z; y, s, u)}{\partial z_i} \tag{2}$$

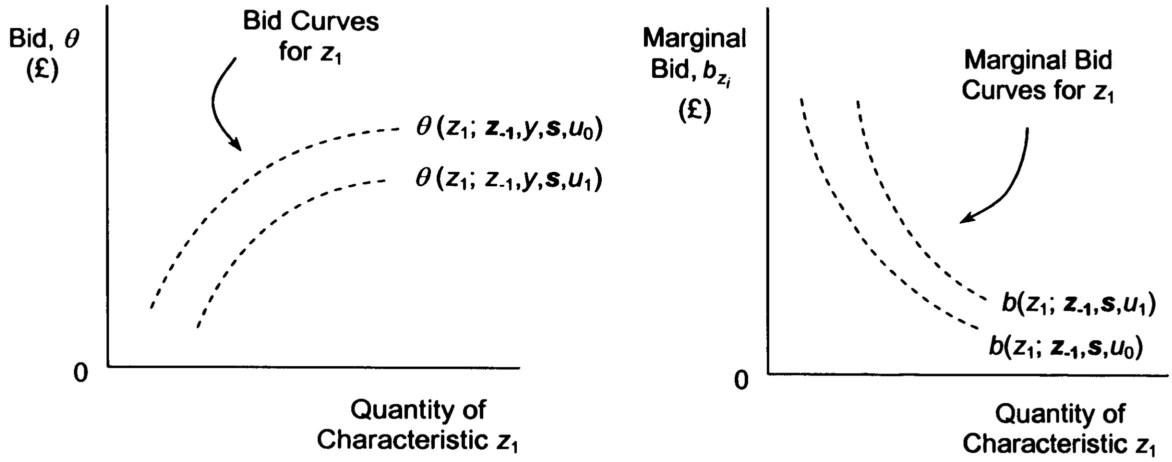
Notice that the household's income y falls out of the marginal bid function. Everything else being equal, the amount that a household is prepared to pay for a property with one extra unit of an attribute in order to maintain the same level of utility is independent of their income.

The marginal bid function can itself be illustrated as a *marginal bid curve* which describes the slope of an equivalent bid curve.

Two bid curves and the equivalent marginal bid curves for a household are illustrated in Figure 1. In the left hand panel, the higher bid curve corresponds to combinations of payments and housing attribute z_1 that result in a utility level u_0. The lower bid curve corresponds to a higher level of utility, u_1, since each level of attribute z_1 is associated with a lower payment.

As we would expect, the marginal bid curves in the right hand panel of Figure 1 slope down from left to right. The household is prepared to pay less for each successive unit of attribute z_1. Though not shown in the figure, at some level of z_1 the marginal bid curves will intercept the horizontal axis. This intercept would reflect the point of satiation at which the household gains no added benefit from purchasing more z_1.

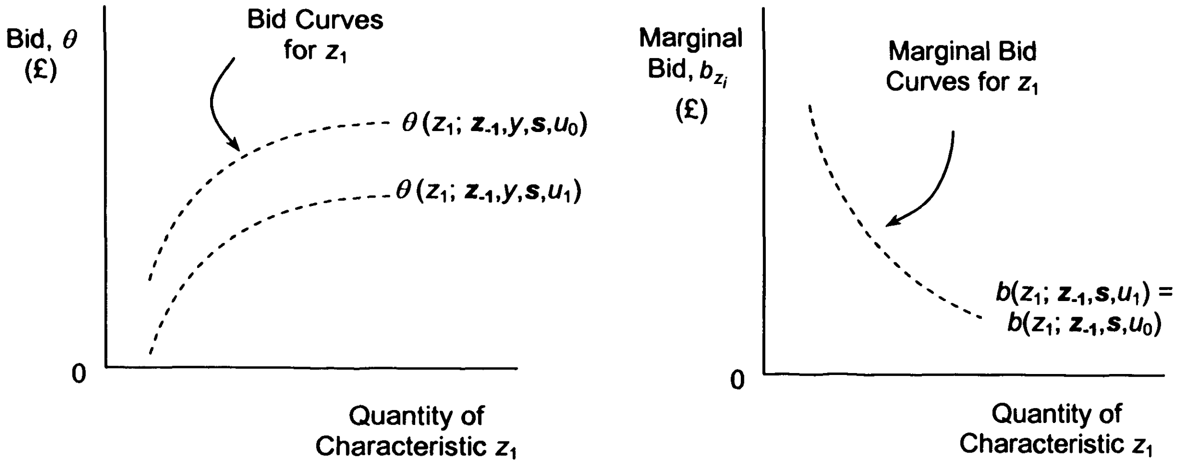
Figure 1: Bid Curves and Marginal Bid Curves



One special case of which we should be aware is when households have quasilinear preferences. This is the case shown in Figure 2. Quasilinear preferences describe indifference curves which are simply vertical translations of each other. Since bid curves are inverted indifference curves, quasilinear preferences can be illustrated as in the left panel of Figure 2 where the bid curves are just vertical translations of each other. Notice that in this case, the slope of the bid curve at all levels of z_1, is identical for all bid curves no matter what level of utility they represent. With quasilinear preferences, therefore, the household's marginal bid functions lie on top of one another. The relevance of this particular form of preferences will become apparent later.

In Chapter 1 we showed how the household's choice of property characteristics could be illustrated using bid functions and the HPF. As shown in the left hand panel of Figure 3, the household chooses the bundle of housing attributes that positions them on the bid curve providing the highest level of utility whilst still being compatible with reigning market prices. In other words, the household maximises their utility by moving to the lowest bid curve that is just tangent with the HPF. In the illustration the household's optimal choice is to select a property with $\hat{z}_1$ of housing attribute z_1. (Notice that we use a hat to signify optimal choices). This property provides the household with their maximum possible utility, u_1.

Figure 2: Bid Curves and Marginal Bid Curves with Quasilinear preferences

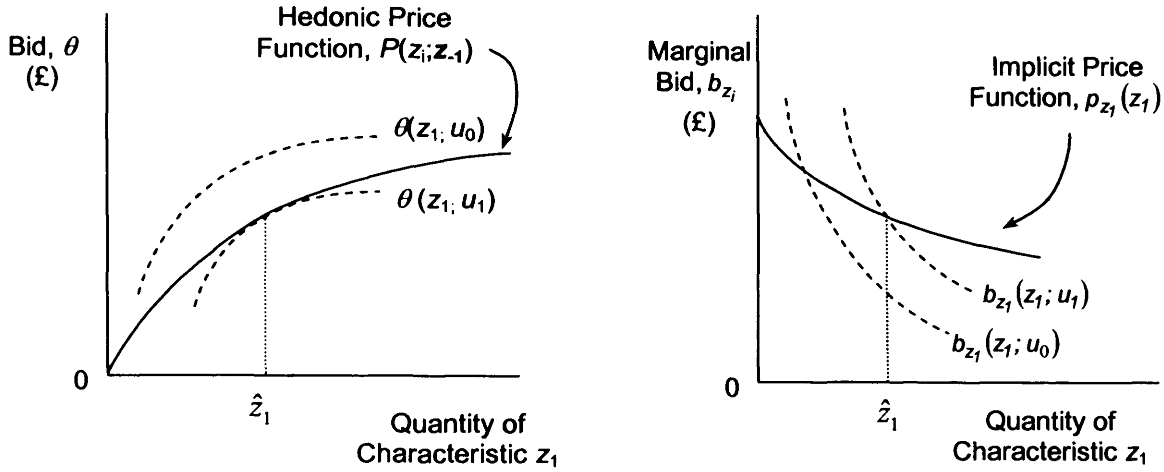


The optimal choice can also be illustrated using marginal bid curves. The right hand panel of Figure 3 plots marginal bid curves corresponding to levels of utility u_0 and u_1. On the same graph is drawn the implicit price function for attribute z_1, $p_{z_1}(z_1)$. Casting our minds back to Chapter 1, remember that the implicit price function describes the additional amount that must be paid by any household in the property market to move to a property with a higher level of characteristic z_1, other things being equal (see Figure 2). The implicit price function is defined mathematically as the derivative of the HPF with respect to attribute z_i. That is;

$$p_{z_i}(z_i; z_{-i}) = \frac{\partial P(z)}{\partial z_i} \tag{3}$$

Thus $p_{z_1}(z_1)$ is the function giving the marginal price of extra z_1. Notice that the implicit price is a function and depends on the level of z_1. (Of course it may also depend on the levels of other housing attributes, z_{-1}, but for simplicity we have suppressed these arguments.) As emphasised in Chapter 1 and illustrated in Figure 3, the implicit price of an attribute does not have to be constant for all levels of z_1.

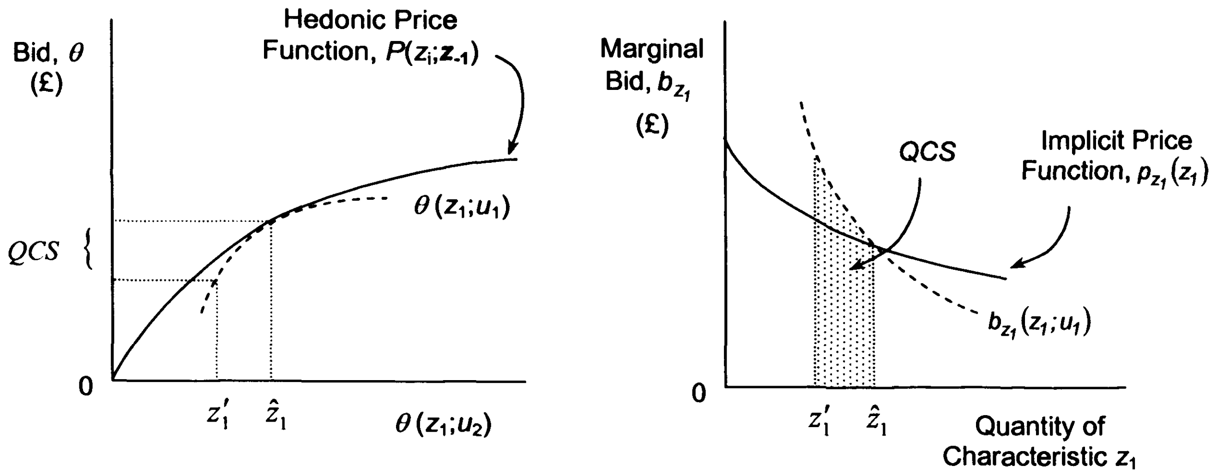
Figure 3: Choice of Optimal Attribute Levels using Bid Functions and Marginal Bid Functions



To establish the choice of attribute levels in the marginal analysis one must know in advance the maximised level of utility, u_1. Then the optimal bundle can be found by moving down the marginal bid curve corresponding to u_1 until the household's marginal WTP for extra z_1 is identical to the marginal price of z_1 in the market[43]. This is very intuitive. The household will always wish to purchase properties with up to $\hat{z}_1$ units of the attribute since their WTP for each of these units is greater than the price of those units. Conversely, the household would not wish to purchase a property with more of attribute z_1 than $\hat{z}_1$, since the price that must be paid for each unit of z_1 in excess of $\hat{z}_1$ is greater than the household's WTP for those units. The optimal level of z_1, therefore, will be found at the intersection of the marginal bid function corresponding to maximised utility and the implicit price function.

[43] In some presentations of hedonic theory, it is not made clear that except for the case of quasilinear preferences, there are an infinite number of marginal bid curves each corresponding to a different level of utility. Moreover, to define the household's optimal choice of housing attributes using marginal bid curves, one must know which of these marginal bid curves corresponds to the maximising level of utility.

Figure 4: Welfare Analysis using Bid Functions and Marginal Bid Functions



The quantity compensating surplus (*QCS*) defined in Chapter 2 can also be illustrated using marginal bid functions. Imagine a household whose optimal residential location has a level of attribute z_1 given by $\hat{z}_1$. An exogenous change decreases the level of z_1 enjoyed at this location to z_1'. The *QCS* measure of welfare change is defined as the amount of money that if given to the household whilst living in the same property would make them as well off as they had been previous to the change. In other words, the household's minimum willingness to accept compensation for suffering the fall in the level of z_1. In the left hand panel of Figure 4 this is illustrated as the difference between the optimising bid curve at $\hat{z}_1$ and z_1'.

Now, since, the marginal bid curve is simply the derivative of the bid curve, this amount is exactly equivalent to the shaded area in the right hand panel of Figure 4. That is, the *QCS* can be measured as the area under the marginal bid curve (corresponding to maximum utility) between the two levels of attribute z_1.

3. Identification of the Marginal Bid Function in Multiple Markets

For a moment, let us consider the problem faced by a researcher investigating a hedonic market. To undertake the project, the researcher collects together information on the selling prices of properties in a single market and records details of the attributes of the property and the characteristics of the purchasing household. Using the data on property prices and attributes, the researcher uses multiple regression techniques to estimate the HPF. This is often referred to as the *first stage* of hedonic analysis.

However, the researcher's objective is to estimate *QCS* measures of welfare changes brought about by changes in the environmental attributes of properties. To estimate such welfare measures the researcher needs to know more than the shape of the HPF. As we have seen, *QCS* measures can be defined in terms of the *bid function* or the *marginal bid function*. Consequently, the researcher must undertake further analysis to estimate either of these two functions. This is often referred to as the *second stage* of hedonic analysis.

Theory tells the researcher that at the optimal choice of attributes the slope of the bid function (corresponding to maximised utility) is equal to the slope of the HPF. Thus, second stage analysis proceeds through the researcher calculating the slope of the HPF at each households choice of property attributes[44].

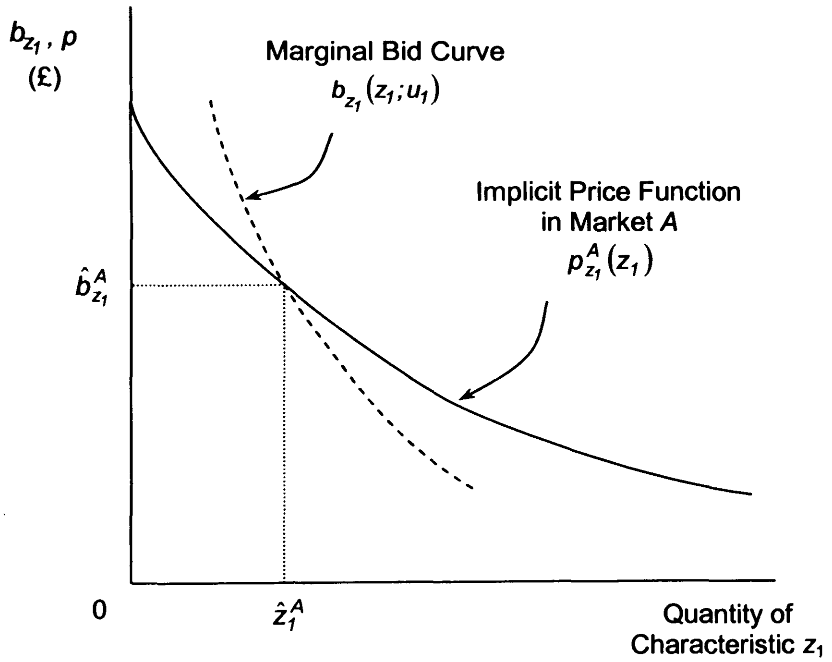
Of course, the slope of the HPF is simply the implicit price of each housing attribute (see Equation 3). Further, as discussed in the previous section, the household's optimal choice of residential location will be such that they equate the implicit price of each housing attribute with the marginal bid curve corresponding to maximised utility (see Figure 3). In short, implicit prices calculated from the first stage analysis provide information on the marginal bid curve. Second stage hedonic analysis, therefore, generally seeks to use the information provided by implicit prices to estimate the marginal bid function.

---

[44] Of course, the slope of the HPF will be multi-dimensional, having as many dimensions as there are housing attributes. In other words, the slope of the HPF, evaluated at any particular combination of property attributes, will describe the implicit price of an extra unit of each housing attribute.

Consider Figure 5. Here the household choosing a property in Market A is faced by the implicit price function for attribute z_1 labelled $p_{z_1}^A(z_1)$. The household chooses a residential location that maximises their utility at level u_1 which corresponds to the marginal bid function shown in the figure. Observing this behaviour in the market, the researcher records just one point on the marginal bid curve. That is, the household's behaviour reveals that for a property boasting $\hat{z}_1^A$ of attribute z_1 the household will be willing to pay $\hat{b}_{z_1}^A$ per unit of z_1 in order to achieve a level of utility u_1.

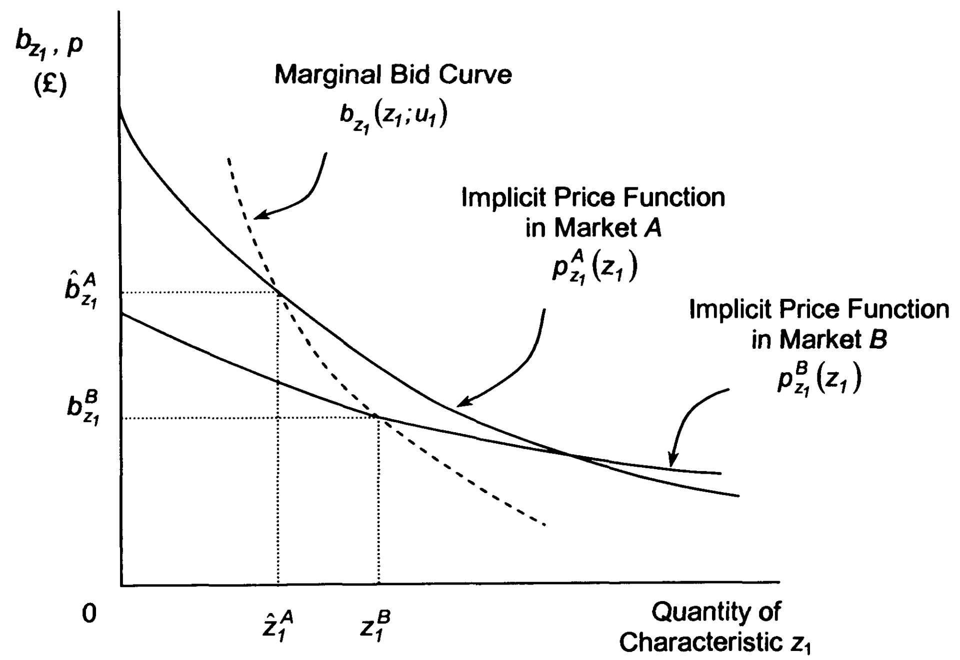
Figure 5: Identifying the Marginal Bid Curve



Unfortunately, knowing one point on the marginal bid curve for u_1 is not sufficient to define the whole curve. Indeed, as various authors have pointed out (e.g. Brown and Rosen, 1982; Murray, 1982; McConnell and Phipps, 1987) any shaped curve is compatible with this one point provided it passes through ($\hat{z}_1^A$, $\hat{b}_{z_1}^A$). To identify the marginal bid function we would require further information. Specifically, we would need to know the household's marginal bids at alternative levels of z_1 that kept the household at a level of utility u_1.

One possibility is that such information could be provided by observing the behaviour of another household in a separate market, market B. If this household happens to have identical income (y) and socioeconomic characteristics (s) to the household choosing in market A, then it is assumed that they will have the same preferences. Thus, if both households faced the same hedonic price schedule they would choose the exact same bundle of attribute levels in their optimal residential location. However, differences in the conditions of supply and demand in the two markets would almost certainly ensure that the equilibrium HPF in market B was different from that in market A.

Figure 6: Identifying the Marginal Bid Curve



This is illustrated in Figure 6 where the non-linear implicit price function for market B, $p^B_{z_1}(z_1)$ is also shown. Notice the second implicit price function cuts the marginal bid function for $b_{z_1}(u_1)$ at ($z_1^B, b_{z_1}^B$). If this were the bundle chosen by the household in market B, then we would have information on the shape of the marginal bid function. Indeed if we could observe the intersection of $b_{z_1}(u_1)$ with a number of different implicit price functions then we would have the required information to trace out the shape of the marginal bid function.

Unfortunately, this is not the case. Since the HPF is different in the second market, the second household's optimal choice of residential location may not afford the same level of utility. For example, if prices are generally lower, then the household's maximised level of utility might also be greater, say u_2. What the researcher would observe in the second market would be the intersection of $b_{z_1}(u_2)$ with $p^B_{z_1}$, and no information would be gained on the shape of $b_{z_1}(u_1)$[45].

We shall return to discuss this predicament in more detail shortly. For now, however, we can draw the following conclusions;

- In order to estimate the marginal bid function, researchers require information on the choices made by similar households faced by different implicit prices. Estimation of marginal bid curves, therefore, requires data from multiple markets.
- The observed behaviour of households' choices in different markets does not provide the information needed to directly estimate the marginal bid function.

4. Marginal Bid functions and Demand Curves with Linear Hedonic Price Functions

Chapter 1 highlighted the fact that households are unable to "repackage" the different attributes of a property. In other words, households cannot break up a property into its constituent parts and enjoy the benefits of each characteristic separate from the whole. It was shown that one of the consequences of this feature of hedonic markets is that the HPF may not be linear. That is, it is possible for the price that is paid for each extra unit of a particular housing attribute to vary according to the level of that attribute. Indeed, typically the additional amount paid for properties enjoying increasingly higher quantities of a characteristic (the implicit price of that characteristic) declines as the total level of that characteristic increases. In this

[45] Unless of course $b_{z_1}(u_2)$ and $b_{z_1}(u_1)$ were identical. This will only happen in the special case where households have *quasilinear* preferences.

section, we return to the issue of non-constant implicit prices and show why this causes problems in the second stage of hedonic analysis.

To illustrate the problem, it is easiest to begin in the counterfactual and assume, for the time being, that implicit prices are constant. Figure 7 depicts the choices made by three identical households[46] selecting a property in three different markets (markets *A*, *B* and *C*). To simplify the problem further, we shall study only one dimension of the households' choice problem; their selection of a level of housing attribute z_1.

Let us focus for the moment, on the choice made by the household in Market *A*. Here the household faces the HPF P^A. Notice that the HPF is *linear* such that the implicit price of z_1 in market *A*, is simply the constant $p_{z_1}^A$.[47] To emphasise this point, when the HPF is linear, the implicit price function can be described by just one parameter, in this case the constant $p_{z_1}^A$.

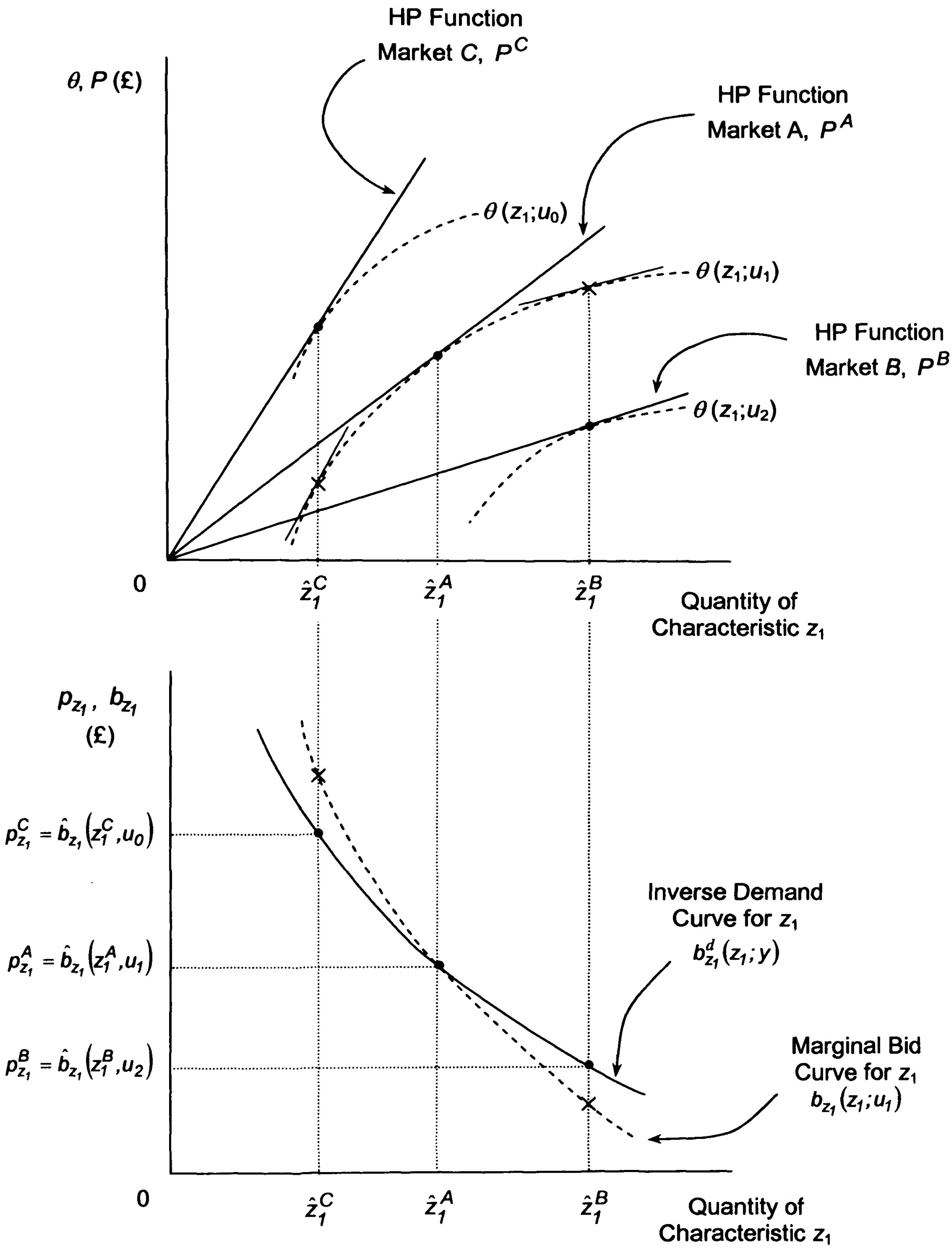
The household in market *A* maximises their utility by moving to the lowest bid curve that is just tangent with the HPF, $\theta(u_1)$. In the illustration the household's optimal choice is to select a property with $\hat{z}_1^A$ of housing attribute z_1. This property provides the household with their maximum possible utility, u_1. This choice point is marked with a dot (as are all other actual choices made by households in the following discussion).We can trace this choice of z_1 down into the lower panel of Figure 7 which shows a marginal analysis of the same information. As discussed in the previous section, the household's marginal bid is given by the implicit price of z_1 at a level of $\hat{z}_1^A$. Since, the HPF is linear the implicit price is simply the constant $p_{z_1}^A$. Hence we can plot one point on the marginal bid curve, $b_{z_1}(z_1;u_1)$, at $\left(\hat{z}_1^A, p_{z_1}^A\right)$.

---

[46] That is, each household has the same income, *y*, and socioeconomic characteristics, *s*. Since the households are identical, we could alternatively treat them as the same household choosing a property in three different markets. Further, since *y* and *s* are identical, these arguments are suppressed in the bid functions and marginal bid functions presented in the text and figures.

[47] Notice that the implicit price is no longer shown as the function $p_{z_1}(z_1)$, where z_1 in brackets indicates that the implicit price depends on the level of z_1.

Figure 7: Linear HPF and inverse demand curves



Now let us turn to the household in market B. Notice that the linear HPF in market B, P^B, has a shallower slope than that in market A. Consequently, the constant implicit price of z_1, $p_{z_1}^B$, in this market is itself lower. Of course, if the price of each unit of z_1 is lower, the household will be able to reach a higher level of overall utility. Indeed, as illustrated in the top panel of Figure 7, the household maximises utility by choosing $\hat{z}_1^B$ of housing attribute z_1. At this choice point the household is on their highest bid curve consistent with the HPF, $\theta(z_1;u_2)$, where they realise the higher level of utility u_2. Again we can plot this choice point on the lower panel at $(\hat{z}_1^B, p_{z_1}^B)$.

Notice, however, that $(\hat{z}_1^B, p_{z_1}^B)$ is not a point on the marginal bid curve $b_{z_1}(z_1;u_1)$.[48] As suggested in the last Section, observing the household's choice of z_1 in a second market with a different implicit price does not provide the researcher with the information necessary to trace out the marginal bid curve $b_{z_1}(z_1;u_1)$. Nevertheless, in our diagrammatic presentation we can locate the point on $b_{z_1}(z_1;u_1)$ corresponding to $\hat{z}_1^B$. The implicit price in market B, $p_{z_1}^B$, is the household's observed WTP for extra z_1 at $\hat{z}_1^B$. The amount we are looking for, however, is the household's marginal WTP for extra z_1 at $\hat{z}_1^B$ whilst maintaining utility u_1.

On the diagram this corresponds to the slope of the bid function $\theta(z_1;u_1)$ at $\hat{z}_1^B$. This point is marked by a cross on the diagram through which a line tangential to the bid function has been drawn. (In the following discussion crosses indicate behaviour not actually observed in markets). Notice that the slope at this point is slightly shallower than that of the HPF in market B. Consequently, the marginal bid curve $b_{z_1}(z_1;u_1)$ at $\hat{z}_1^B$ will itself be slightly lower than the observed marginal bid at $\hat{z}_1^B$ (i.e. $p_{z_1}^B$). This point is marked on the lower diagram in Figure 7 with a cross.

48 Rather it is a point on the marginal bid curve $b_{z_1}(z_1;u_2)$. Again, the marginal bid curve $b_{z_1}(z_1;u_2)$ will be different to $b_{z_1}(z_1;u_1)$ unless the household has quasilinear preferences.

In general, this will be the case for any attribute if it behaves like a normal good. Only if the household has quasilinear preferences will the two slopes be identical at $\hat{z}_1^B$. If this were the case the dot and cross in the lower diagram would coincide.

Finally, observe the choice made by the household in market *C*. Here the implicit price of z_1 is the constant $p_{z_1}^C$. Since this is higher than that observed in either of the other markets, the household in market *C* must make do with a lower level of utility. Indeed, the utility maximising choice of z_1, $\hat{z}_1^C$, only affords a level of utility u_0. Again we can plot the observed behaviour in the lower panel as the point $\left(\hat{z}_1^C, p_{z_1}^C\right)$. Meanwhile, the point corresponding to $\hat{z}_1^C$ on the marginal bid curve $b_{z_1}(z_1;u_1)$ is the slope of $\theta(z_1;u_1)$ at $\hat{z}_1^C$. Notice that this is slightly steeper than the HPF in market *C*. Hence the marginal bid for z_1 that maintains the level of utility u_1 is higher than the marginal bid observed in the market $p_{z_1}^C$. This point is also plotted in the lower panel of Figure 7. Again if preferences were quasilinear then the dot and cross would coincide.

So far we have managed to plot five points in the lower panel of Figure 7. Those marked with dots represent choices actually observed in the market, those marked with crosses represent behaviour not actually observed.

In fact these five points trace out two separate curves. The first, constructed by joining the dots, is what we would actually observe if we were to collect data on household's property choices from different markets with linear hedonic functions. This curve traces out household's marginal WTP for extra z_1 at different levels of z_1. For those familiar with economics, this is simply an *inverse ordinary demand curve*. We denote this function;

$$b_{z_1}^d(z_1;y) \tag{4}$$

Where;

- $b_{z_1}^d(\cdot)$ is the inverse ordinary demand function for housing attribute z_1

- z_1 is the level of the housing attribute and
- y is the household's income

With a linear HPF, the inverse ordinary demand function takes a very simple form sloping down from left to right. As we might expect, at higher levels of z_1 the household is willing to pay less for each extra unit.

The second curve is that which the researcher wishes to identify, the marginal bid curve. This traces out household's marginal bids at different levels of z_1 that maintain a level of utility u_1. For those familiar with economics, this is simply an *inverse compensated demand curve*. As already stated, we denote this function;

$$b_{z_1}(z_1; u) \tag{2}$$

Where;

- $b_{z_1}(\cdot)$ is the marginal bid curve or inverse compensated demand function for housing attribute z_1
- z_1 is the level of the housing attribute and
- u is the level of utility

Unfortunately, this second curve is not observed in market behaviour. Crucially, however, the inverse ordinary demand curve and the marginal bid curve will generally be fairly similar (as shown pictorially in the figure).

Indeed, they will be identical if the household has quasilinear preferences. Quasilinear preferences represent the special case where the household has a zero income elasticity of demand for the housing attribute. Remember from Equation (1) that increases in income translate directly (i.e. pound for pound) into increases in the bid function. In effect, increases in income cause, the bid curves to shift vertically upwards. Since quasilinear preferences give rise to bid curves that are themselves vertical translations of each other the net effect of an increase in income is that the household moves onto a bid curve representing a higher level of utility but does not change their demand for the good.

In the real world, however, quasilinear preferences are the exception rather than the rule. One might reasonably expect that as a household's income increases their demand for housing attributes would itself increase. Moreover, the greater the income elasticity of demand for the particular attribute the greater the difference between the ordinary inverse demand curve and the marginal bid curve.[49] However, under reasonable assumptions concerning the income elasticity of demand, theoretical research indicates that the difference between these two curves is unlikely to be sufficiently large to warrant the extra complexity of deriving the marginal bid curve (Willig, 1976; Randall and Stoll,1980).

One possibility, therefore, is that researchers use market data to estimate the ordinary inverse demand curve. Approximate QCS welfare measures can be estimated as the area under the inverse demand curve between the two levels of attribute z_1. Further, if this approximation is thought to result in serious error, there are techniques by which the researcher can retrieve the marginal bid curve from an estimated inverse demand curve, we shall return to this in later discussion.

5. Marginal Bid Functions and Demand Curves with Nonlinear Hedonic Price Functions

In a world with purely linear HPFs, therefore, everything appears rosy. Market data can be used to estimate the inverse demand function and this should provide a reasonably good approximation to the marginal bid function. However, in the real world, HPFs are not linear and there is the rub. When implicit prices are not constant and preferences are not quasilinear, the inverse demand curve as we have illustrated it does not exist.

To illustrate observe Figure 8. Here we have done away with the assumption of linear HPFs and quasilinear preferences. Now the HPFs in markets A, B and C are all non-linear. The figure has been constructed such that the households in the three markets maximise their utility by choosing the exact same levels of z_1 as were

[49] The difference between the slopes of the two curves will also depend on the significance of expenditure on that attribute as a part of the consumer's budget.

illustrated in the linear case of Figure 7. Further, the diagram has been drawn such that the household in market A achieves the same level of utility, u_1, at their optimal choice of z_1 as was chosen facing the linear HPF in Figure 7.

By construction, therefore, the point in the lower panel of Figure 8 corresponding to the choice of the household in market A, is identical to that in Figure 7; $(\hat{z}_1^A, p_{z_1}^A)$. Once again, this describes one point on the marginal bid function $b_{z_1}(z_1;u_1)$.

Consider now the choice of the household in market B. The non-linear HPF in this market is in all places lower than that in market A. Consequently, the price paid for any level of z_1 in market B is less than that paid for the same level of z_1 in market A. Not surprisingly, therefore, the household in market B, manages to achieve a higher level of utility, u_3, whilst choosing a higher level of z_1, $\hat{z}_1^B$.

Following a now familiar procedure, we can plot this choice point in the lower panel of Figure 8 by determining the implicit price of z_1 at $\hat{z}_1^B$ as the slope of the bid function $\theta(z_1;u_2)$ at $\hat{z}_1^B$. Notice that with a non-linear HPF, the implicit price at $\hat{z}_1^B$ may not be the same as the implicit price at other levels of z_1.

In the linear case, this choice point defined a second point on the inverse ordinary demand curve. Indeed, we might expect that in this non-linear case we could trace out a similar shaped curve. Certainly this second point in the lower panel of Figure 8 would seem to be following the correct pattern. As we would expect, the household's WTP for z_1 at this higher level of provision is lower than that observed at the lower level of provision chosen in market A. Further, if we plot the marginal bid function $b_{z_1}(z_1;u_1)$ at this level of provision it falls below that observed in market choices. Again the result observed in the linear HPF case.

However, observe the choice made by the household in market C. Since the HPF is in all places higher than that in market A, it comes as no surprise that the household's optimal choice, is at a lower level of provision and affords them a lower level of overall utility, u_{-1}. When we come to plot this choice point in the lower panel, however, we are struck by an anomaly. At $\hat{z}_1^C$ facing the HPF in market C, the household's marginal WTP for extra z_1 is lower than that recorded in market A. This is despite the fact that the household in market C has chosen a property with lower levels of z_1 than that chosen in market A.

Clearly, with non-linear HPFs and preferences that are not quasilinear, observed choices do not plot out a nice downward sloping inverse ordinary demand curve[50]. To emphasise this point consider Figure 9 where a fourth identical household is observed choosing a property in market D. Here, the household maximises their utility by choosing $\hat{z}_1^D$ of the housing attribute. Whilst this is an identical quantity to that chosen by the household in market A, the slope of the HPF in market D is shallower than that in market A. Plotting this on the marginal analysis diagram we see that with nonlinear HPFs, the same level of demand can be associated with two different implicit prices. To summarise, when implicit prices are non-constant and preferences are not quasilinear, the inverse ordinary demand curve as normally conceived is not well defined. A household's marginal WTP for extra z_1 at any level of z_1 will depend on the shape of the entire HPF faced in that market.

The problem is further complicated when we move out of the unidimensional problem of choosing just one housing attribute level and consider choice across many attributes. In this case, if patterns of substitutability and complementarity exist between the attribute of interest and the other attributes, then the household's marginal WTP for extra z_1 at any level of z_1 will also depend on the shape of the HPF for all these attributes.

[50] Note that if preferences were quasilinear then the slope of the bid function at any particular level of housing attribute would be the same for all bid curves. In this special case, the existence of nonlinear HPFs does confound the existence of a downward sloping inverse ordinary demand curve.

Figure 8: Non-linear HPF and inverse demand curves; Example 1

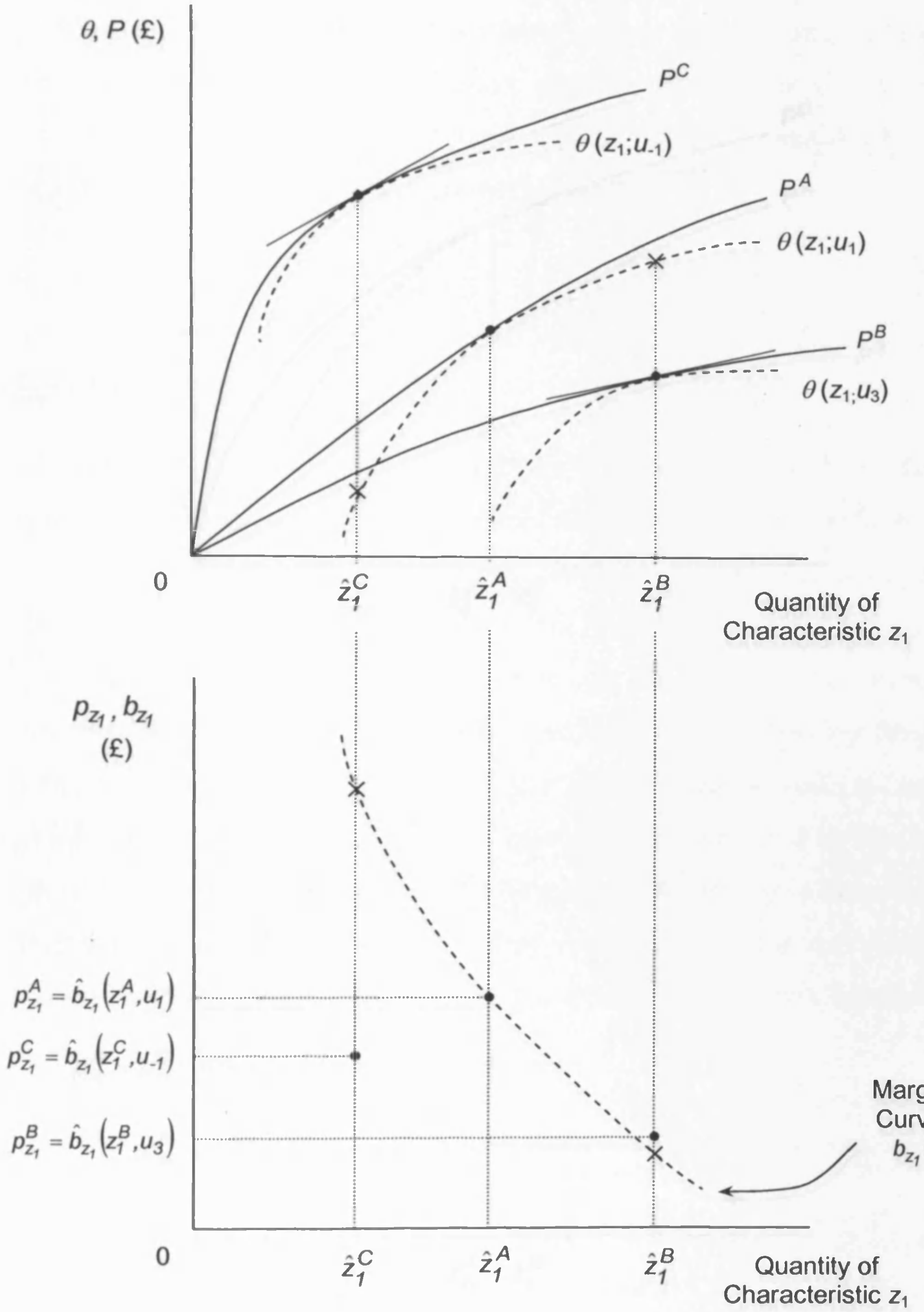
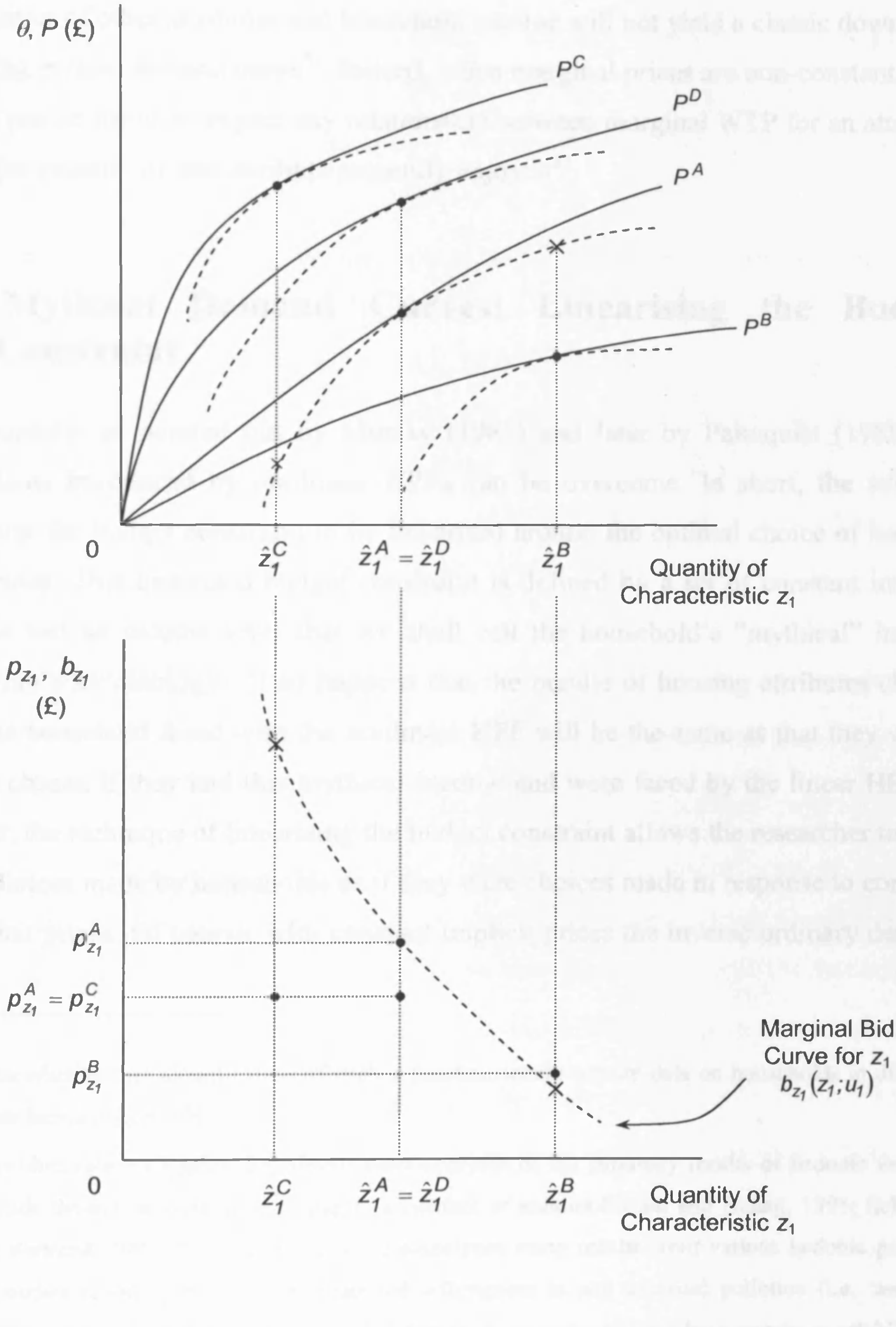


Figure 9: Non-linear HPF and inverse demand curves; Example 2



This presents a considerable problem for welfare analysis in hedonic markets. Specifically, it becomes impossible to estimate a simple inverse ordinary demand function for an attribute of interest. That is, a simple regression of the implicit prices paid for an attribute by different households against quantities of this attribute, quantities of other attributes and household income will not yield a classic downward sloping inverse demand curve[51]. Indeed, when marginal prices are non-constant there is no reason for us to expect any relationship between marginal WTP for an attribute and the quantity of that attribute presently enjoyed[52].

6. Mythical Demand Curves: Linearising the Budget Constraint

Fortunately, as pointed out by Murray (1983) and later by Palmquist (1988) the problems introduced by nonlinear HPFs can be overcome. In short, the solution requires the budget constraint to be linearised around the optimal choice of housing attributes. This linearised budget constraint is defined by a set of constant implicit prices and an income level that we shall call the household's "mythical" income (Murray's terminology). It so happens that the bundle of housing attributes chosen by the household faced with the nonlinear HPF will be the same as that they would have chosen if they had this mythical income and were faced by the linear HPF. In effect, the technique of linearising the budget constraint allows the researcher to treat the choices made by households as if they were choices made in response to constant implicit prices. Of course, with constant implicit prices the inverse ordinary demand

[51] Remembering that identification of such a function would require data on households in different markets facing different HPFs

[52] This observation suggests that simple meta-analyses of the summary results of hedonic analyses have little theoretical basis. For example, a number of authors (Smith and Huang, 1995; Schipper, 1996; Bertrand, 1997) have carried out meta-analyses using results from various hedonic property price studies reporting households' marginal willingness to pay to avoid pollution (i.e. 'average' implicit prices for pollution). Amongst other things, these meta-analyses have sought to establish the relationship between marginal willingness to pay to avoid pollution and current levels of pollution. The discussion in this section shows that in the face of non-linear HPFs, no simple relationship between the two exists.

function is defined by Equation (4) and takes on its classic downward sloping curve. This "mythical" inverse ordinary demand function should be a reasonable approximation to the household's marginal bid curve.

The technique of linearising the budget constraint is illustrated in Figure 10. The top panel of this diagram depicts the choice of housing attribute z_1 made by two households faced by the same nonlinear HPF. Let us assume that these two households have the same socioeconomic characteristics, *s*, but that household *b* has a higher income than household *a*. That is y_b is greater than y_a.

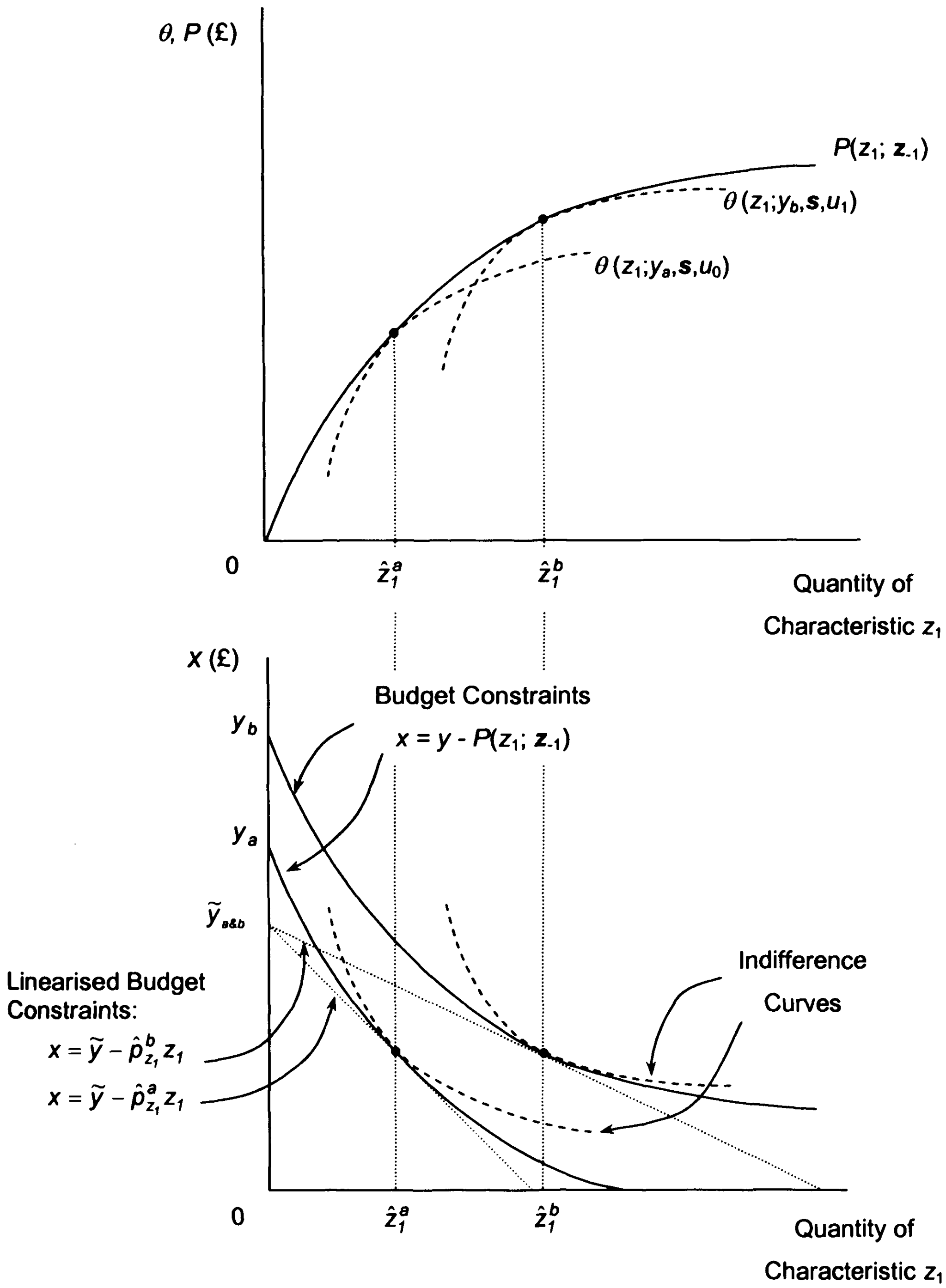
We can just as well illustrate these choices in the indifference diagram in the lower panel. This diagram plots indifference relationships between money to spend on other goods, the numeraire, and the level of housing attribute z_1. Since the HPF is nonlinear, the budget constraints faced by the two households are themselves nonlinear. Notice that the budget constraint for household *b* is simply a vertical translation of that faced by household *a*. The actual incomes of the two households will be given by the point where the budget constraints intercept the y-axis and these two amounts are labelled on the diagram[53].

Consider now the choice made by household *a*. This household optimises their utility by choosing a level of the housing attribute labelled $\hat{z}_1^a$ at which the implicit price of z_1 is $\hat{p}_{z_1}^a$. At this point we wander into the realms of the "mythical" rather than real worlds. Imagine that the implicit price at this optimal choice of housing attributes was actually a constant marginal price coming from a linear hedonic function. If this were so we could construct a budget constraint running through the household's optimal choice with a slope of $\hat{p}_{z_1}^a$. The intercept of this mythical budget constraint gives household *a*'s mythical income $\tilde{y}_a$. The important thing to note is that the choice of property attributes made by household a with income y_a facing the

[53] We assume that households would not be willing to pay anything for a house with no z_1. For example, if z_1 represents "peace and quiet", then this assumption amounts to saying that there is a point where a household would not purchase a property because it is too noisy to live in.

nonlinear HPF is identical to that which they would have made if they had an income of $\tilde{y}_a$ and faced a linear hedonic function with constant marginal price $\hat{p}^a_{z_1}$.

Figure 10: Linearising the budget constraint



Now consider the choice made by household b. Following the same procedure, we can construct a mythical linear budget constraint whose slope is defined by the implicit price of the attribute at household b's optimal choice, $\hat{p}^b_{z_1}$. The intercept of this budget constraint with the vertical axis gives household b's mythical income $\tilde{y}_b$. Again, the bundle of attribute quantities chosen by household b will be identical whether they are making choices in the real world with the nonlinear hedonic function and income y_b or in the mythical world with the linear HPF and income $\tilde{y}_b$.

The diagram has been constructed such that both households have the same mythical income. Notice that the decisions made by these two households could just as well be treated as those made by a single mythical household with income $\tilde{y}_{a\&b}$ choosing a property in two separate markets. In the first market this mythical household faces a linear HPF in which z_1 has the constant implicit price $\hat{p}^a_{z_1}$ in the second the household faces a linear HPF with the slightly lower constant implicit price $\hat{p}^b_{z_1}$. As we would expect, the household facing the lower price chooses more z_1. Indeed, given observations from many households with the same mythical income we could trace out the entire mythical ordinary demand curve. Since in the mythical world all HPFs are linear, the mythical ordinary demand curve is well defined.

The critical question for welfare analysis is the extent to which mythical demand curves approximate compensated demand curves. Once again, this crucially depends on income effects. If there are no income effects then the relationship between prices and quantities is not influenced by adjustments to expenditure. In this case, the mythical demand curve will exactly replicate the compensated demand curve. As income effects become more pronounced, the mythical demand curve will begin to differ from the compensated demand curve in ways that mirror the divergence between the ordinary demand curve and the compensated demand curve in a world of constant marginal prices.

Theoretical work by Edlefsen (1981) and Blomquist (1989) has sought to describe the relationships that exist between the mythical and ordinary demand functions when marginal prices are not constant. Unfortunately, there appears to be no theoretical research comparing the mythical and compensated demand curves in a world of non-constant marginal prices to parallel that carried out by Willig (1976), Randall and Stoll (1980) and Hanemann (1991) in comparing the ordinary and compensated demand functions in a world of constant marginal prices. The author acknowledges that this is a weakness of the current study.

7. Mythical Demand Curves: Estimation and welfare analysis

The linearisation procedure described in the last section suggests a reasonably simple procedure by which data on property prices might be used to recover household preferences for environmental goods traded implicitly in the property market. In the first place, the HPF must be accurately estimated for several markets or independent market segments and implicit prices recovered for each housing attribute. Subsequently, the mythical incomes of the households purchasing the properties in the data set can be calculated according to:

$$\tilde{y} = y - P(\hat{\mathbf{z}}) + \sum_{i=1}^{K} \hat{p}_i \hat{z}_i \qquad (5)$$

The mythical inverse ordinary demand curve can then be estimated by regressing the implicit price of each attribute on the chosen quantities of that attribute, the chosen quantities of other attributes and mythical income;

$$\hat{p}_{z_i} = \tilde{b}_{z_i}\left(\hat{z}_i, \hat{\mathbf{z}}_{-i}, \tilde{y}, \mathbf{s}\right) \qquad (i = 1 \text{ to } K) \qquad (6)$$

More typically, researchers estimate the mythical ordinary demand function;

$$\hat{z}_i = \tilde{z}_{z_i}\left(\hat{p}_{z_i}, \hat{\mathbf{p}}_{\mathbf{z}_{-i}}, \tilde{y}, \mathbf{s}\right) \qquad (i = 1 \text{ to } K) \qquad (7)$$

where $\hat{\mathbf{p}}_{\mathbf{z}_{-i}}$ is the vector of all other attribute chosen implicit prices.

Equation (7) tends to be seen as a more natural specification than Equation (6) since it is the z_i rather than the p_{z_i} which are the observed outcome of household's choices in hedonic markets. Note carefully, however, that in hedonic markets, where marginal prices are nonlinear household's actually simultaneously choose both the quantities and the marginal price of attributes.

As indicated in Equation (7), in an ideal world, the researcher would estimate a system of demand curves for all property attributes. In reality, however, the usual procedure is to concentrate on one or a number of attributes that form the focus of the research programme. Further, rather than including all attribute quantities in the regression and imposing the theoretical restrictions on Equations (6) and (7) required by demand theory, researchers employ fairly simple functional forms, including only a handful of other attribute quantities.

Econometric estimation of mythical ordinary demand curves is further complicated by problems of endogeneity. As we have seen, in hedonic markets, the marginal price of housing attributes will generally not be constant. In maximising their utility from the choice of residential location, the household chooses both the quantity of housing attributes and the marginal price of the attributes. In estimating, Equation (7), therefore, the implicit prices of housing attributes on the right hand side of the equation are *endogenous*. Further, since mythical income is calculated using the chosen level of marginal price (Equation 5), this too is endogenous. Unless researchers account for this endogeneity, the parameter estimates from the econometric estimation of the mythical inverse ordinary demand curve will be biased.

Typically, endogeneity is handled through the application of instrumental variable techniques. The trick here is to regress each of the endogenous variables in the demand equation on a set of exogenous variables that in this context are referred to as instruments. The results of these ancillary regressions are used to calculate predicted values for the endogenous variables. The demand equations are estimated using these predicted rather than the actual values of the endogenous variables. Avoiding the econometric details, the instrumental variables technique removes the problem of biased parameter estimates caused by the inclusion of endogenous regressors in the demand equations.

This all seems very straightforward, however, difficulties arise in choosing suitable instruments. These variables should be highly correlated with the endogenous variable they are being used to predict but at the same time should not be correlated with the error term entering the demand equation. For example, imagine that we were choosing instruments for the household's mythical income. Suitable candidates might include the household's socioeconomic characteristics including the number of members of the household, their ages and educational status. Suitable instruments for implicit prices could once again include socioeconomic traits but authors have also suggested using the marginal price paid by similar households, where similarity is determined either in terms of these household's socioeconomic characteristics (Murray, 1983) or their spatial proximity (Cheshire and Sheppard, 1998).

With the mythical ordinary demand curve estimated, approximate *QCS* measures of welfare change can be obtained by integrating under this curve between the initial level of the attribute and that following some external change.

Table 1 presents a step by step guide to hedonic analysis, from collecting data through to welfare estimation.

Table 1: Steps to Perform a Hedonic Analysis

<table>
<tr><td>Step 1</td><td>Collect data

This should include;

• Property sales prices and
• the socioeconomic characteristics of purchasing households

Data should provide information on the choices made by households in two or more independent hedonic property markets.</td></tr>
<tr><td>Step 2</td><td>Estimate Hedonic Price Function for each market

Regress property prices on property characteristics according to;

$$P = P(z_1, z_2, \ldots, z_K)$$

• Repeat for each separately identified property market
• Test for market segmentation with each property market</td></tr>
<tr><td>Step 3</td><td>Calculate Implicit Prices chosen by Households

For each household, calculate the implicit price of housing attributes according to;

$$p_{z_i}(z_i; z_{-i}) = \frac{\partial P(z)}{\partial z_i}$$</td></tr>
</table>

Step 4	*Calculate each Household's Mythical Income* Using the implicit prices estimated in step 3 calculate each household's mythical income according to; $$\tilde{y} = y - P(\hat{z}) + \sum_{i=1}^{K} \hat{p}_i \hat{z}_i$$
Step 5	*Calculate instruments for Implicit Prices and Mythical Income* Select instruments for implicit prices. Suitable candidates include; • Socioeconomic characteristics • Implicit prices chosen by similar (e.g. in demographic traits and/or spatial proximity) households Select instruments for Mythical Income. Suitable candidates include; • Socioeconomic characteristics Using data from all markets estimate two ancillary equations regressing observed implicit prices and mythical income on instruments Use the regression results to calculate predicted values for implicit prices and mythical income for each household. Call these; $\breve{\tilde{y}}$ and $\breve{p}_{zi}$
Step 6	*Estimate Mythical Ordinary Demand Function* Using predicted values calculated in step 5 estimate the demand function according to; $$\hat{z}_1 = z_{z_1}^{M}\left(\breve{p}_{z_1}, \breve{\boldsymbol{p}}_{\boldsymbol{z}_{-1}}, \breve{\tilde{y}}, \boldsymbol{s}\right)$$
Step 7	*Calculate QCS welfare measures* Integrate under the mythical demand curve between the initial level of the attribute and that following some external change

8. Mythical Demand Curves: Benefits Transfer

Whilst the techniques of demand estimation from hedonic analysis have been known for some years, the majority of empirical applications have stopped short of estimating mythical demand curves. Rather researchers have gone no further than Step 3, estimating the HPF and reporting the implicit price of housing attributes. Whilst implicit prices can be used for measuring the welfare impacts of marginal changes in housing attributes in a particular market, they will not be accurate indicators of the welfare impacts for large changes in the housing attribute or when changes occur over a wide geographic area (see discussion in Chapter 2). Further, these implicit prices are specific to a particular housing market since they are determined by the particular circumstances of supply and demand operating in that market. Consequently, there is no theoretical basis for transferring implicit prices from one market to another. Benefits transfer using implicit prices is meaningless.

Recently, a number of research articles have reported more thorough hedonic analyses in which demand curves have been estimated (e.g. Cheshire and Sheppard, 1998; Palmquist and Isangkura, 1999; Boyle et al., 1999 and Zabel and Kiel, 2000). These studies seek to identify underlying household preferences for housing attributes. Consequently they can be used to derive theoretically consistent estimates of welfare changes. Importantly, under the assumption that household preferences for housing attributes are stable across different property markets, such demand functions should be transferable across property markets. In the next Chapter, we apply the techniques described in this Chapter to estimate demand functions for the City of Birmingham data set.

CHAPTER 9: WELFARE ANALYSIS FOR THE CITY OF BIRMINGHAM PROPERTY MARKET

1. Introduction

In Chapter 8 a method by which household preferences could be estimated using data from households' choices of residential location in property markets was proposed. The key relationship identified by this method is the mythical ordinary demand function, given by;

$$\hat{z}_i = \tilde{z}_{z_i}\left(p_{z_i}, \boldsymbol{p}_{z_{-i}}, \tilde{y}, \boldsymbol{s}\right) \qquad (i = 1, 2, ..., K) \qquad (1)$$

where, $\tilde{z}_{z_i}(\cdot)$ is the mythical demand function for property attribute z_i, $\hat{z}_i$ is the quantity of that attribute chosen by a household, given the attributes implicit price, p_{z_i}, the vector of the other attributes' implicit prices, $\boldsymbol{p}_{z_{-i}}$, the socioeconomic characteristics of the household, $\boldsymbol{s}$, and the household's mythical income, $\tilde{y}$, defined as;

$$\tilde{y} = y - P(\hat{z}) + \sum_{i=1}^{K} \hat{p}_i \hat{z}_i \qquad (2)$$

In this Chapter we build on the results of Chapter 5 to estimate mythical ordinary demand functions for the City of Birmingham dataset. We estimate separate demand curves for exposure to road traffic noise and exposure to rail traffic noise. These are used to provide estimates of the welfare benefits of reducing these two types of noise pollution. The results described in this Chapter are the main policy-relevant findings of the research presented in this thesis.

2. Data Issues

To move from the general form of Equation (1) to a tractable expression that may be estimated econometrically we face a number of challenges.

First, observe that the demand system defined in Equation (1) implies that the demand for each property attribute is a function of the marginal prices of all other attributes.[54] Unfortunately the large number of property attributes contained in the HPF would imply the inclusion of an impractically large number of explanatory variables. To simplify, we focus solely on the estimation of demand functions for the noise pollution attributes of a property, and include in the demand system only the own-price and cross-price terms for these attributes.

Second, the Birmingham dataset does not contain details of household income. Consequently, it is impossible to calculate mythical income as defined in Equation (2). Rather, we replace income, y, in Equation (2) with total expenditure on property services; that is, with the full price of the property purchased by the household. Such an assumption can be viewed in one of two ways;

- Either we take a puritanically theoretical viewpoint and invoke a two-stage budgeting hypothesis. In this case, households are assumed to first decide how much they are prepared to spend on property and second how to allocate this expenditure between property attributes (given the prices of those attributes implicit in the hedonic price schedule) to arrive at their optimal choice of residential location.
- Alternatively, we accept that we do not have the theoretically correct variable to include in our regression but believe that property expenditure will act as a reasonably accurate proxy for the missing income data.

As we shall see later, the latter interpretation may be of use in facilitating benefits transfer. For now, however, we remain true to the theoretical framework.

Given these assumptions, let us redefine the general model described in the previous section. Households are assumed to choose a residential location so as to maximise the *subutility* function pertaining to property choice, which we denote;

$$U^{sub}(z, x; s) \tag{3}$$

Here the arguments are defined as,

---

[54] Indeed, we have already considerably simplified the demand relationship by assuming that all non-property goods can be taken as a composite good with a price of one (i.e. all other goods are condensed into the one good "money to spend on other things").

- z, a K-vector ($K = 2$) of noise pollution levels deriving from two sources; road and rail traffic.
- x, the numeraire which now represents expenditure on all other property attributes.
- s, the characteristics of the household.

Analysis is made easier if we continue to assume that the property attribute levels in the subutility function are goods. Consequently, we invert the scale for measuring levels of noise, such that our measure indicates the level of "peace and quiet".

Given the two-stage budgeting assumption, the subutility function for property attributes is maximised with respect to the budget allocated to property expenditure. In our application, therefore, the budget constraint is taken as;

$$y = P(z, x) \tag{4}$$

that is, we redefine y to be a household's total expenditure on property.

Moreover, by subsuming all other expenditure on property characteristics into the numeraire attribute x we simplify calculation of mythical income by linearising property expenditure (Equation 4) with respect to only the noise pollution variables;

$$\begin{aligned} \tilde{y} &= P(\hat{z}, \hat{x}) - \left(P(\hat{z}, \hat{x}) - P(z = 0, \hat{x})\right) + \sum_{i=1}^{K} \hat{p}_i \hat{z}_i \\ &= P(z = 0, \hat{x}) + \sum_{i=1}^{K} \hat{p}_i \hat{z}_i \end{aligned} \tag{5}$$

The demand system in (1) also requires data on the characteristics of households. Unfortunately, this data is also absent from the data set. As a proxy we include the factor scores describing the socioeconomic characteristics of the household's chosen neighbourhood. Since the wealth factor is found to be highly collinear with the total expenditure variable we drop this argument from the demand equation contending that total property expenditure is a better proxy for household wealth.

The final data issue concerns the fact that our data relates to property purchases. For the purposes of cost-benefit analysis what is required is the value of changes in noise pollution per period of time. Unfortunately, the prices we have estimated in the first stage of the hedonic analysis are equivalent to prices paid for durable goods; one pays "up-front", then enjoys the benefit of that purchase over the lifetime of the product. Rather, we would prefer to have our prices expressed per period, say annually. Of course, in the property market such annual payments are known as property rents.

If markets are operating perfectly, and generally we assume that they are, then the price at which the household purchases the property will be the discounted sum of all the future per period rents from the property according to;

$$P(z,x)=\sum_{t=1}^{T}\frac{R(z,x)}{(1+d)^{t}} \tag{6}$$

where $R(z,x)$ is the equivalent per period rent on a property with market price $P(z,x)$, t indexes each time period, T is the expected life of the property and d is the discount rate. For Equation (6) to provide an expression that can be used to convert prices into annual rents we must first decide upon values for T and d. It is mathematically convenient (and not too unrealistic) to assume that the expected life of the property, T, can be taken as infinite. In this case (6) simplifies considerably and provides us with the following straightforward expression relating prices to rents;

$$R(z,x)= d\cdot P(z,x) \tag{7}$$

The same expression can be used to change implicit prices for property characteristics into implicit rents;

$$r_{z_i} = d\cdot p_{z_i} \tag{8}$$

and total expenditure on property services into an annualised expenditure on property services;

$$m = d\cdot y \tag{9}$$

which, given the assumption of Equation (4), is simply Equation (7) using different notation. Of course, for our purposes we are concerned with the "mythical" (linearised) version of (9) given by;

$$\tilde{m} = d \cdot \tilde{y} \tag{10}$$

The rate of discount, *d*, in this relationship should reflect the rate at which households are prepared to trade off present with future expenditure.[55] In the context of property purchases a number of rates suggest themselves (see Table 1).

Imagine that a household purchasing a property were able to provide all the capital for the purchase from their own savings. In such a case they would forego the interest payments that they would receive on that capital over time. As such, we might assume that in a well-functioning property market the amount they would be prepared to pay in rent each year would be the foregone interest. If rents were cheaper than this the household would prefer to rent, if they were more expensive then it would make sense to purchase the property. If this were true of all households then the laws of arbitrage suggest that *d* would settle at the rate of interest that could be earned by depositing money in a bank or building society.

Alternatively, a household may have to borrow money from a lender in order to purchase a property. In such a case, they would be expected to pay back that loan over time at the mortgage interest rate. If rents were less than these payments then it would make sense to rent rather than purchase a property. Alternatively if interest repayments were less than rents then rational households would choose to take out a mortgage and purchase a property. Again, if this were true of all households then arbitrage would suggest the *d* would settle at the mortgage interest rate.

Unfortunately, the real world is far more complex. Households finance property purchases from a variety of sources, usually comprising a combination of personal savings and borrowings. Table 1 provides details from the Office of National Statistics "Financial Statistics Time Series" of some key interest rates in 1997 (the year from which the Birmingham data is taken) and for the present day. The second

---

[55] And, as such, includes an element reflecting the household's expectations concerning the devaluing worth of money through price inflation.

column provides details of the interest rate on savings, the final column provides the mortgage rate and the third column lists the base rate of interest in the economy. To assess which of these rates most closely resembles the value taken by d, we propose examining data for current property prices and rents in Birmingham. Clearly, it would be preferable to use data comparing 1997 prices and rents but no such data were available to the author.

Table 1: Interest rates in 2003 and 1997

	Time Deposit Rate[1]	Base Rate[2]	Mortagage Rate[3]
1997 (Avg)	5.75	6.50	7.07
2003 (Feb)	2.32	3.75	4.93

Sources:

[1] ONS series THAN: Weighted avg of bank & building society interest rates on time deposit accounts

[2] Bank of England base rates

[3] ONS series AJNL: Weighted avg of building society mortgage rates

Numerous estate agent and letting agent listings were examined in order to find pairs of very similar properties that were being offered for rent and sale respectively. Properties were matched on a variety of features, namely location (to first 4 digits of the postcode), number of bedrooms, property type and construction age. Table 2 provides details of the small sample collected for this purpose.

Using Equation (7) the implied values of d are listed in the final column of Table 2. To garner a best estimate of d from this data a no constant OLS regression of Equation (7) was executed. This provided an estimate of d of 4.39% with 95% confidence intervals ranging from 4.01% to 4.77%. Note that the simple model of Equation (7) provides a good fit to this data (adj R^2 value .962) lending support to the theoretical basis for this relationship between prices and rents.

Compare this confidence interval to the range of rates provided for 2003 in Table 1. Clearly, the interest rate on bank savings underestimates the value of d. Conversely, the average mortgage rate provides an overestimate. The value of d suggested by the 2003 data lies somewhere between these two extremes but closer to the mortgage rate. Such a pattern may be expected if the funds for financing property purchases come more from mortgage borrowing than from savings. Indeed a simple calculation shows that for 2003 the implied value of d lies almost exactly halfway between the mortgage rate and the base rate.

Table 2: Prices and Rents for matched properties in Birmingham (March, 2003)

Obs	Bedrooms	Age	Type	Location	Price (£)	Rent/Year (£)	Implied value of d
1	3	Between Wars	Semi-detached	Sutton Coldfield	194,950	8,220	4.22%
2	3	Victorian	Terrace	Erdington	129,995	7,020	5.40%
3	3	Between Wars	Semi-detached	Sutton Coldfield	159,950	7,020	4.39%
4	2	Victorian	Terrace	Harborne	190,000	7,500	3.95%
5	3	Post War	Semi-detached	Kings Norton	143,500	6,300	4.39%
6	4	Between Wars	Semi-detached	Sparkhill	199,950	7,800	3.90%
7	3	Between Wars	Semi-detached	Stechford	129,950	5,100	3.92%
8	3	Between Wars	Semi-detached	Harborne	197,950	7,500	3.79%
9	3	1970s	Terrace	Harborne	185,000	6,300	3.41%
10	2	Victorian	Terrace	Moseley	77,950	5,400	6.93%
11	3	Between Wars	Semi-detached	Selley Oak	129,950	6,900	5.31%
12	2	Victorian	Terrace	Northfield	105,000	5,100	4.86%
13	2	Victorian	Terrace	Kings Norton	92,500	4,800	5.19%
14	3	Modern	Detached	Northfield	157,950	8,700	5.51%
15	4	Modern	Detached	Hall Green	209,950	11,400	5.43%
16	3	Post War	Semi-detached	Hall Green	95,000	6,000	6.32%
17	2	Post War	Terrace	Kings Heath	81,950	5,400	6.59%
18	2	Victorian	Terrace	Balsall Heath	85,000	5,400	6.35%
19	4	Edwardian	Semi-detached	Harborne	350,000	14,400	4.11%
20	3	Between Wars	Semi-detached	Harborne	249,950	7,800	3.12%
21	3	Modern	Semi-detached	Harborne	185,000	9,000	4.86%
22	3	Modern	Semi-detached	Acocks Green	155,000	8,340	5.38%

We use this as a simple rule of thumb and calculate the value of d for 1997 to be halfway between the 1997 mortgage rate and the 1997 base rate; a value of 6.78%. In employing this rule of thumb, we must assume a certain degree of myopia on behalf of house buyers. That is, households are assumed to calculate d based purely on current interest rates and not on the expected future values of these rates as would be more theoretically correct. In reality, this assumption is probably not too far-fetched. For a start, there is little reason to assume that households can predict interest rate changes over anything but the immediate future and certainly not over the lifetime of a property. Furthermore, rather than considering the purchase price of a property, the relevant quantity to the average home buyer is their prospective annual outlay on mortgage payments and their foregone returns on savings. In reality, it is this value that households compare to rental prices when deciding whether to buy or rent. Let us assume that the rental market is sufficiently liquid to allow rental prices to fall in line with the average prospective per period payments of home buyers and thereby restore equilibrium to the property market. In this case, there is no need to interpret d as a discount factor, rather the value of d simply records the average ratio of annual payments to purchase price. Since homebuyers are usually tied into mortgage deals for no more than a few years and are unlikely to consider long term changes in interest rates in their decisions it seems reasonable to assume that the value of d will be intimately related to the interest rates prevailing in the economy at the time of purchase. In the absence of evidence to the contrary we assume that the relationship between d and these interest rates does not change as the interest rates themselves change.

Finally, notice that the first stage hedonic regression returns results that allow for negative prices for peace and quiet (most notably in submarkets one and two). Clearly, negative prices defy both theory and common sense; are we to believe that households are prepared to pay more for properties in noisier locations? As a result we treat these negative prices as the result of sampling error where the 'true' implicit prices for peace and quiet are in the region of zero. Accordingly we set these prices to a value of zero in the regression analysis. Details of the regressor data used in the second stage analysis are given in Table 3.

Table 3: Means and percentiles of the regressor data for the 2nd stage analysis

Statistics	Total Prices			Annual Prices			Socioeconomic composition of households in Neighbourhood		
	Road Price	Rail Price	Property Price	Road Price	Rail Price	Property Rent	Ethnic	Age	Family
Mean:	149.6	369.4	59,710	10.14	25.04	3,551	0.04	-0.14	0.03
Percentiles:									
5th	0	0	26,500	0	0	1,764	-0.91	-1.69	-1.45
10th	0	0	31,000	0	0	2,093	-0.79	-1.33	-1.08
25th	76.9	170.1	39,000	5.22	11.53	2,644	-0.58	-0.67	-0.54
50th	134.3	297.9	50,000	9.10	20.20	3,390	-0.30	-0.05	0.06
75th	228.9	563.2	67,950	15.52	38.18	4,558	0.19	0.47	0.63
90th	302.5	849.0	96,000	20.51	57.56	6,505	1.67	0.88	1.11
95th	350.9	1,003.3	127,000	23.79	68.02	8,539	2.59	1.15	1.46

3. Econometric Issues

Our econometric analysis seeks to estimate two separate demand equations; one for peace and quiet from road traffic noise a second for peace and quiet from rail traffic noise. In general form, we express these two relationships as;

$$\begin{aligned} q_1^* &= \tilde{q}_1(r_1, r_2, \tilde{m}, s) \\ q_2^* &= \tilde{q}_2(r_1, r_2, \tilde{m}, s) \end{aligned} \qquad (11)$$

which is simply an annualised version of the (mythical) demand system of Equation (1) for these two particular goods. We denote equations and variables pertaining to road noise by the subscript 1 and those pertaining to rail noise by the subscript 2. Furthermore, q_k^* (k = 1, 2) is the quantity of peace and quiet; $\tilde{q}_k(\cdot)$ (k = 1, 2) is the annualised mythical demand relationship, r_k (k = 1, 2) are the annualised implicit prices for peace and quiet, $\tilde{m}$ is linearised annual expenditure on property services and s is a vector of socioeconomic demand shifters.

To maintain simplicity we estimate (11) as a linear demand system[56] according to;

$$q^{*}_{ki} = \boldsymbol{Y}_i \boldsymbol{\beta}_k + \boldsymbol{X}_{1i} \boldsymbol{\gamma}_k + u_{ki} = \boldsymbol{Z}_i \boldsymbol{\delta}_k + u_{ki} \quad (k = 1,2;\ i = 1,\ldots,N) \tag{12}$$

Here we have gathered together the implicit price (r_1 and r_2) and expenditure ($\tilde{m}$) observations for household i into the 1×3 vector $\boldsymbol{Y}_i$. Moreover, for reasons that will become clear shortly, we have redesignated the vector of socioeconomic demand shifters (s) for household i as the vector $\boldsymbol{X}_{1i}$. In our application the socioeconomic variables are proxied by neighbourhood factors describing the ethnicity, age composition and family composition of the local population. The explanatory variables of the demand function for household i can be gathered together into the vector $\boldsymbol{Z}_i = [\boldsymbol{Y}_i, \boldsymbol{X}_{1i}]$ whilst the parameters of the demand functions can be gathered together in the vector $\boldsymbol{\delta}_k = [\boldsymbol{\beta}_k, \boldsymbol{\gamma}_k]$ $(k = 1,2)$. Finally, u_{ki} is an unobservable household specific random term capturing elements of the demand relationships not accounted for in the model specification.

3.i. Corner Solutions: Censored Regression

Numerous problems confound the estimation of the parameters of Equation (12) through the straightforward application of familiar regression techniques such as ordinary least squares (OLS).

[56] In matter or fact, economic theory implies numerous relationships between the variables and parameters of the demand system that are contravened by the linear specification of Equation (12). Several econometric specifications have been developed that seek to explicitly impose the restrictions implied by economic theory. One example of such a specification is the almost ideal demand system (AIDS) of Deaton and Muellbauer (1980), a specification which was employed by Sheppard and Cheshire (1998) in their demand analysis based on hedonic data. Here we forego the niceties of imposing theoretical restrictions. Rather, for numerous reasons, we prefer the linear specification of Equation (12). In particular, in analysing our data we need to overcome a number of econometric "difficulties", most notably censoring of the dependent variable and the endogeneity of some of the regressors. Techniques for dealing with these difficulties are well developed for linear regression. Furthermore, the linear functional form facilitates ease of interpretation of the parameter estimates and simplifies calculation of welfare estimates.

Specifically, observe that in our dataset containing 10,641 observations, only 2,723 households choose locations enduring road traffic noise above the urban background level (taken as 55dB), whilst only 379 choose locations enduring rail traffic noise above the urban background level. Furthermore, we assume that households are unconcerned by traffic noise levels that fall below the urban background noise. As such we transform the noise measure in three ways;

- First, we subtract the urban background noise level (55db) from the measured noise levels at each property. As such, households choosing properties with positive noise measures will be in locations in which traffic noise exceeds the background noise, whilst those choosing properties with negative measures will be in locations unaffected by traffic noise pollution.
- Second, we convert our measure of the "bad", noise pollution, into the "good", peace and quiet, by multiplying through by –1. As such, increasing values of our measure now indicate increasingly peaceful locations. This is our version of the q_k^* measure introduced in the previous section.
- Finally, we set all values of the noise measure that are positive to zero. That is, we assume that households will be indifferent to increasing peace and quiet from traffic sources that exceed the background level of peace and quiet to be found in urban areas.

Let us denote the transformed noise variable by q_k. Note that it is q_k (when measured as a bad) not q_k^*, that is used in the first stage estimation of implicit prices. Our dependent variable, peace and quiet from traffic sources, can be described as follows;

$$q_{ki} = \begin{cases} q_{ki}^* & if\ q_{ki}^* < 0 \\ 0 & otherwise \end{cases} \qquad \left(k = 1,2; \quad i = 1,2,\ldots,N\right) \tag{13}$$

In econometric parlance, our dependent variable is censored from above.

Unfortunately, in the presence of censoring, standard approaches to the estimation of a linear function such as (13) will return biased and inconsistent estimates of the model parameters. An alternative estimator is offered by the Tobit model. The Tobit model is defined by the combination of Equations (12) and (13). Moreover, we impose the assumption that the error term u_{ki} is distributed normally with mean zero and constant variance σ_k^2. Furthermore, we assume that the error terms are uncorrelated across households and, for now, that the error terms are uncorrelated across demand relationships even for the same household. That is;

$$\begin{aligned} & u_{ki} \sim N\left(0, \sigma_k^2\right) \\ & \operatorname{cov}\left(u_{ki}, u_{kj}\right) = 0 \quad \left(i \neq j\right) \qquad \left(i, j = 1, 2, \ldots, N; k = 1, 2\right) \\ & \operatorname{cov}\left(u_{1i}, u_{2j}\right) = 0 \end{aligned} \tag{14}$$

Under these assumptions, it is possible to establish the distribution of the censored variable q_{ki} and estimate the parameters of the model through maximum likelihood techniques. We apply these techniques and report the parameter estimates in Table 4.

Whilst the Tobit model adequately handles censoring of the dependent variable, there are other econometric issues that must be addressed in estimating the parameters of the model. As such, we defer a comprehensive discussion of the parameter estimates to a later date.

Here we simply note that, as would be expected, the own-price terms in both demand equations are negative and significant; the higher the price of peace and quiet the less peace and quiet is demanded by households. Likewise the cross-price terms are negative indicating *complementarity* between peace and quiet from different sources. That is to say, the households' objective is to achieve a quiet environment. To do this they must purchase not only peace and quiet from road traffic, but also peace and quiet from rail traffic; purchasing one without the other is of little benefit. Finally, in both demand equations the expenditure variable has a coefficient taking the expected positive sign and is a highly significant determinant of the level of demand for peace and quiet.

Table 4: Road and rail noise demand parameters from application of the Tobit Model

	Road		Rail	
	Coefficient (Std. Err.)	***t*-score**	**Coefficient (Std. Err.)**	***t*-score**
Road Noise Price (r_1)	-0.180*** *(0.031)*	-5.75	-0.224*** *(0.062)*	-3.64
Rail Noise Price (r_2)	-0.007 *(0.010)*	-0.964	-0.043** *(0.022)*	-1.98
Log of Linearised Expenditure ($\tilde{m}$)	3.361*** *(0.525)*	6.32	8.136*** *(1.087)*	7.49
Ethnicity Factor	-0.612*** *(0.184)*	3.33	0.719* *(0.372)*	1.93
Age Factor	-0.135 *(0.212)*	-0.64	1.621*** *(0.409)*	3.96
Family Factor	-1.271*** *(0.203)*	-6.25	0.664* *(0.397)*	1.67
Constant	-16.697*** *(4.056)*	-4.12	-36.384*** *(8.061)*	-4.51
σ^2	13.499*** *(0.219)*		14.354*** *(0.662)*	
Log Likelihood:	-14,471.2		-2,630.4	
Uncensored Obs:	2,723		379	
Censored Obs:	7,918		10,262	

*** Significant at 99% level of confidence
** Significant at 95% level of confidence
* Significant at 90% level of confidence

3.ii. Endogenous Variables: Instrumental Variables Regression

A further problem hindering the straightforward estimation of the demand relationships is one of endogeneity. As discussed in detail in Chapter 1, implicit prices in a hedonic market are non-parametric; that is, the marginal price is not constant for all units of quality characteristics. In such a market, households' choices are not constrained to choosing the quantity of a characteristic given the constant market price for each unit. Rather households actually simultaneously choose both the quantities and the marginal price of attributes.

In such a situation, marginal prices are not exogenous to the choice problem. Indeed, since prices are chosen by the households they must be treated as endogenous. Unfortunately, in the presence of endogeneity, econometric theory predicts that the parameter estimates from the Tobit model will be biased and inconsistent (for a more detailed exposition see, for example, Davidson and MacKinnon; 1993).

In our case, the problem is further compounded by the fact that the cross-price terms must also be considered as endogenous to the choice problem. And, what is more, since linearised expenditure ($\widetilde{m}$) is calculated using these same marginal prices, this must also be considered as an endogenous explanatory variable.

The fact that the own-price, cross-price and expenditure terms must be treated as endogenous explains why these three variables have been gathered together into vector $\boldsymbol{Y}_i$, distinct from the variables assumed to be exogenous that are gathered together into the vector $\boldsymbol{X}_{1i}$.

The most general technique for handling endogenous variables is the method of *instrumental variables* (IV). The fundamental ingredient of any IV procedure is a set of instrumental variables. Crucially, each instrumental variable must be correlated with the endogenous variables in the model, but must be independent of those elements of the joint decision process (i.e. the simultaneous choice of quantity and price) that are not captured by the model. Here we assume that these uncaptured elements reflect specific characteristics of the household and their attitude to noise. As such we follow the suggestion of Cheshire and Sheppard (1998) and select as instrumental variables the values of r_{jk} and $\widetilde{m}_j$ chosen by other nearby households. We define "nearby" as proximity in a multidimensional space defined by geographic location, socioeconomic characteristics (i.e. neighbourhood factors) and property characteristics (i.e floor and garden area). Our matrix of instrumental variables consists of r_{jk} and $\ln\left(\widetilde{m}_j\right)$ for the three nearest neighbours to each observation and the squares of these values. We gather the instrumental variables for each household into the vector of exogenous regressors $\boldsymbol{X}_{2i}$.

If we wished to be explicit, we could write out the econometric equations describing households' choice of these endogenous variables. Such equations are termed *structural equations* by econometricians since they describe the underlying structure of the choice problem being described. By substituting such equations into one another, it is possible to reformulate structural equations so that only exogenous variables appear on the right hand side. Such equations are termed *reduced-form* equations. We shall describe the process of producing a reduced-form equation in a short while. For now, however, observe that the reduced-form equations for the endogenous explanatory variables can be written as the set of three linear equations;

$$\boldsymbol{Y}_i = \boldsymbol{X}_i \boldsymbol{\Pi} + \boldsymbol{v}_i = \boldsymbol{X}_{1i} \boldsymbol{\Pi}_1 + \boldsymbol{X}_{2i} \boldsymbol{\Pi}_2 + \boldsymbol{v}_i \qquad (i = 1, \ldots, N) \tag{15}$$

Here, $\boldsymbol{Y}_i$ is each household's observed choice of road price, rail price and expenditure; $\boldsymbol{X}_{1i}$ is the vector of exogenous variables from the demand equations (the socioeconomic factors) and $\boldsymbol{X}_{2i}$ is the vector of instrumental variables describing chosen prices and expenditures of similar households. These two vectors of exogenous variables are gathered together into the vector $\boldsymbol{X}_i$. The parameters of the reduced form equations for the endogenous variables are given by the matrices $\boldsymbol{\Pi}_1$ and $\boldsymbol{\Pi}_2$, which can be gathered together into the matrix $\boldsymbol{\Pi}$. Finally, $\boldsymbol{v}_i$ is a vector of disturbances.

To construct reduced form demand equations we may substitute (15) into (12) giving;

$$\begin{aligned} q^*_{ki} &= (\boldsymbol{X}_i \boldsymbol{\Pi} + \boldsymbol{v}_i)\boldsymbol{\beta}_k + \boldsymbol{X}_{1i} \boldsymbol{\gamma}_k + \boldsymbol{u}_{ki} \\ &= \boldsymbol{X}_{1i}(\boldsymbol{\Pi}_1 \boldsymbol{\beta}_k + \boldsymbol{\gamma}_k) + \boldsymbol{X}_{2i}(\boldsymbol{\Pi}_2 \boldsymbol{\beta}_k) + \boldsymbol{v}_i \boldsymbol{\beta}_k + \boldsymbol{u}_{ki} \\ &= \boldsymbol{X}_{1i} \boldsymbol{\alpha}_{1k} + \boldsymbol{X}_{2i} \boldsymbol{\alpha}_{2k} + \boldsymbol{v}_i \boldsymbol{\beta}_k + \boldsymbol{u}_{ki} \\ &= \boldsymbol{X}_i \boldsymbol{\alpha}_k + \boldsymbol{v}_i \boldsymbol{\beta}_k + \boldsymbol{u}_{ki} \end{aligned} \qquad (k = 1,2;\ i = 1, \ldots, N) \tag{16}$$

where $\boldsymbol{\alpha}_{1k} = \boldsymbol{\Pi}_1 \boldsymbol{\beta}_k + \boldsymbol{\gamma}_k$, $\boldsymbol{\alpha}_{2k} = \boldsymbol{\Pi}_2 \boldsymbol{\beta}_k$ and $\boldsymbol{\alpha}_k$ gathers together the matrices $\boldsymbol{\alpha}_{1k}$ and $\boldsymbol{\alpha}_{2k}$.

Amemiya (1979) suggested that one could estimate $\hat{\boldsymbol{\Pi}}$, the reduced form parameters in (15), using OLS. Subsequently, it is possible to estimate the reduced form parameters of the demand equations, $\hat{\boldsymbol{\alpha}}_k$, by inserting the estimated residuals, $\hat{\boldsymbol{v}}_i$ into (16) and applying Tobit regression techniques.

Of course, our interest is not in the parameters of the reduced form equations, $\boldsymbol{\alpha}_k$, but in the parameters of the structural demand equations $\boldsymbol{\delta}_k$ $(k = 1,2)$. However, Amemiya (1979) observed that $\boldsymbol{\alpha}_k$ and $\boldsymbol{\delta}_k$ are related through the equation;

$$\boldsymbol{\alpha}_k = D(\boldsymbol{\Pi})\boldsymbol{\delta}_k \qquad (k = 1, 2) \tag{17}$$

where $D(\boldsymbol{\Pi}) \equiv [\boldsymbol{\Pi}, \boldsymbol{I}_1]$ is the selection matrix made up of ones and zeros such that $\boldsymbol{X}_{1i} = \boldsymbol{X}_i\boldsymbol{I}_1$.

If we knew the true values of the reduced form parameters $\boldsymbol{\alpha}_k$ and $\boldsymbol{\Pi}$ then it would be a simple task to retrieve $\boldsymbol{\delta}_k$ using (17). Unfortunately, all we have are our estimates of the reduced form parameters, $\hat{\boldsymbol{\Pi}}$ and $\hat{\boldsymbol{\alpha}}_k$. Since these estimates will not be exactly equal to the true values, Equation (17) can be reformulated as per Amemiya (1979) to give;

$$\begin{aligned}\hat{\boldsymbol{\alpha}}_k &= D(\hat{\boldsymbol{\Pi}})\boldsymbol{\delta}_k + (\hat{\boldsymbol{\alpha}}_k - \boldsymbol{\alpha}_k) + \beta(\hat{\boldsymbol{\Pi}} - \boldsymbol{\Pi}) \\ &= D(\hat{\boldsymbol{\Pi}})\boldsymbol{\delta}_k + \eta_k \qquad\qquad (k = 1, 2)\end{aligned} \tag{18}$$

where $\eta_k = (\hat{\boldsymbol{\alpha}}_k - \boldsymbol{\alpha}_k) + \beta(\hat{\boldsymbol{\Pi}} - \boldsymbol{\Pi})$. Amemiya (1979) shows how $\boldsymbol{\delta}_k$ can be estimated by applying generalised least squares to (18) in which the weighting matrix is constructed from a consistent estimator of the asymptotic covariance matrix of $\sqrt{N}(\hat{\boldsymbol{\alpha}} - D(\hat{\boldsymbol{\Pi}}))$. To construct such a weighting matrix requires assumptions to be made concerning the disturbance terms of the reduced form equations for the endogenous explanatory variables (15) and those in the demand equations of (12). Amemiya assumes that conditional on the exogenous variables $\boldsymbol{X}_i$, these disturbance terms are multivariate normal;

$$(u_{ki}, \boldsymbol{v}_{ki}) \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}_k) \qquad (i = 1,2,\ldots,N;\ k = 1,2) \tag{19}$$

The actual form of the weighting matrix is beyond the scope of this discussion. The formulas can be found in Newey (1987).

This procedure for estimating δ_k has become known as Amemiya Generalised Least Squares (AGLS). Amemiya (1979) and Newey (1987) show that these estimates of the structural parameters are more efficient than other possible estimators including that of Smith and Blundell (1986) that we shall discuss shortly.

Table 5 provides estimates of the structural parameters of the demand equations estimated using AGLS.

	Road		**Rail**	
	Coefficient (Std. Err.)	***t*-score**	**Coefficient (Std. Err.)**	***t*-score**
Road Noise Price (r_1)	-0.419*** *(0.100)*	-4.19	-0.399** *(0.197)*	-2.03
Rail Noise Price (r_2)	0.104 *(0.071)*	1.47	-0.446*** *(0.140)*	-3.18
Log of Linearised Expenditure ($\widetilde{m}$)	7.649*** *(2.614)*	2.93	20.297*** *(5.210)*	3.90
Ethnicity Factor	1.438*** *(0.247)*	5.81	0.729 *(0.501)*	1.46
Age Factor	-1.506*** *(0.430)*	-3.50	2.978*** *(0.840)*	3.55
Family Factor	1.084*** *(0.253)*	4.28	0.535 *(0.497)*	1.08
Constant	-52.735*** *(19.431)*	-2.71	-124.726*** *(38.432)*	-3.25
Uncensored Obs:	2,723		379	
Censored Obs:	7,918		10,262	

*** Significant at 99% level of confidence
** Significant at 95% level of confidence
* Significant at 90% level of confidence

In accordance with expectations, the own-price terms in both demand equations are negative and highly significant. Interestingly, the two own-price coefficients are relatively close in value suggesting the slopes of the two demand curves are relatively similar. We expect the cross-price terms to be negative, reflecting the complementarity relationship that exists between peace and quiet from different sources. The regression results, provide weak evidence supporting this hypothesis. In the rail demand equation the parameter on the road noise price is negative and significant at the 95% level of confidence. However, contrary to expectations, the parameter for rail noise price in the road demand equation is positive though not statistically significant.

Demand for peace and quiet is greater for households allocating greater mythical expenditure to the purchase of their property. As Blomquist (1989) observes, it does not necessarily follow that demand is increasing in *actual* expenditure though in numerous special cases this will also be true. In contrast to the similarity in own-price parameter estimates, the parameter estimate for expenditure is a great deal larger in the rail traffic noise equation. This indicates that the income elasticity of demand for peace and quiet from rail traffic noise is greater than that for road traffic noise. It seems that wealthier households are increasingly more likely to choose environments free from rail traffic noise than from road traffic noise.

The remaining covariates consist of the neighbourhood factor scores. Clearly, these variates are not ideal since they are not necessarily indicative of the socioeconomic characteristics of the households themselves. As such, the interpretations that follow must be regarded with a certain dose of cautious scepticism. Noting this fact, we continue our discussion under the assumption that the factor scores are at least an approximate guide to true household characteristics.

Reassuringly, the parameter estimates on the family composition factor are of the same sign and almost identical magnitude. These indicate that demand for noise avoidance is greater amongst households with young families. The ethnicity factor coefficient also takes the same sign in both equations indicating that households from the ethnic minorities demand relatively greater levels of peace and quiet, all else equal. However, for rail traffic noise at least, the parameter estimate is small in magnitude and statistically insignificant. The coefficient estimates for the adult age composition are the least easy to interpret. The estimates suggest that older households demand relatively more peace and quiet from rail traffic noise but, conversely, relatively less peace and quiet from road traffic noise. Since we have no particular theory to support these conflicting results we suggest that this is an appropriate juncture at which to execute "cautious scepticism".

3.iii. Testing the Econometric Specification

Whilst AGLS provides efficient estimates of the model parameters, it does not provide a natural framework within which to test the model specification. In particular, we are interested in answering two questions about our estimated model;

1. Are the variables in Y_i really endogenous?
2. Do the estimated demand relationships for rail and road traffic noise differ from each other?

Smith and Blundell (1986) propose an alternative to AGLS estimation that will prove useful in answering these questions. They observed that under the normality assumption of Equation (19), the distribution of the disturbance of the structural equations, conditional on the chosen values of Y_i, is given by;

$$u_{ki} \sim N\left(\mathbf{v}_i \boldsymbol{\rho}_k, \sigma_k^2\right) \qquad \left(i = 1,2,\ldots,N;\, k = 1,2\right) \tag{20}$$

where $\boldsymbol{\rho}_k$ is the vector of correlation coefficients given by $\boldsymbol{\rho}_k = \boldsymbol{\Sigma}_{k22}^{-1}\boldsymbol{\Sigma}_{k21}$ and σ_k^2 is the conditional variance of u_{ki} given by $\sigma_k^2 = \Sigma_{k11} - \Sigma_{k12}\Sigma_{k22}^{-1}\Sigma_{k21}$, and the partitioning of $\boldsymbol{\Sigma}_k$ conforms with $\left(u_{ki}, \mathbf{v}_{ki}\right)$. Of course, given the conditional distribution of the disturbance term, the conditional distribution of the structural demand equations (13) follows;

$$q_{ki}^{\bullet} \sim N\left(\mathbf{Z}_i \boldsymbol{\delta}_k + \mathbf{v}_i \boldsymbol{\rho}_k, \sigma_k^2\right) \qquad \left(i = 1,2,\ldots,N;\, k = 1,2\right) \tag{21}$$

Hence, given $\mathbf{v}_i$, one can build the conditional likelihood of the model from which the structural parameters, $\boldsymbol{\delta}_k$, can be estimated.

In particular, Smith and Blundell (1986) propose estimating the reduced form equations for the endogenous explanatory variables, by OLS so as to form estimates of the disturbance terms. Let us denote these $\hat{\mathbf{v}}_i$. Including $\hat{\mathbf{v}}_i$ as regressors in the Tobit regression defined by Equations (21) and (13) provides an alternative approach to recovering the structural parameters $\boldsymbol{\delta}_k$. Smith and Blundell term this estimation procedure a "control function" approach. Parameter estimates from the control function estimation procedure are reported in Table 6.

Observe that the parameter estimates from the control function approach are almost identical to those from AGLS. This is not at all surprising considering both provide consistent estimates of the model parameters.

Table 6: Road and rail noise demand parameters using the control function approach (reporting unadjusted standard errors)

	Road		Rail	
	Coefficient (Std. Err.)	***t*-score**	**Coefficient (Std. Err.)**	***t*-score**
Road Noise Price (r_1)	-0.422*** (*0.099*)	-4.28	-0.404** (*0.192*)	-2.10
Rail Noise Price (r_2)	0.107 (*0.070*)	1.54	-0.451*** (*0.135*)	-3.35
Log of Linearised Expenditure ($\widetilde{m}$)	7.635*** (*2.586*)	2.95	20.732*** (*4.998*)	4.15
Ethnicity Factor	1.448*** (*0.244*)	5.94	0.748 (*0.493*)	1.52
Age Factor	-1.531*** (*0.424*)	-3.61	3.059*** (*0.823*)	3.72
Family Factor	1.088*** (*0.250*)	4.36	0.565 (*0.485*)	1.16
$\hat{v}$ (Road Noise Price)	0.248** (*0.103*)	2.40	0.190 (*0.202*)	0.94
$\hat{v}$ (Rail Noise Price)	-0.119* (*0.071*)	-1.68	0.416*** (*0.136*)	3.07
$\hat{v}$ (Expenditure)	-5.305** (*2.641*)	-2.01	-12.478*** (*5.052*)	-2.47
Constant	-52.602*** (*19.227*)	-2.74	-127.069*** (*36.884*)	-3.45
σ^2	13.478*** (*0.219*)		14.324*** (*0.660*)	
Log Likelihood:	-14,452.2		-2,624.8	
Uncensored Obs:	2,723		379	
Censored Obs:	7,918		10,262	

*** Significant at 99% level of confidence
** Significant at 95% level of confidence
* Significant at 90% level of confidence

Since the disturbance terms included in the control function approach are estimated parameters, Smith and Blundell (1986) show how the standard Tobit model variance matrix should be adjusted. Note that the standard errors reported in Table 6 do not make this adjustment. However, Smith and Blundell (1986) also note that if all the explanatory variables are actually exogenous, then the parameter estimates on the control function variables, ρ_k, will be insignificant and the variance matrix for the estimates of the other parameters will collapse back to that provided by the standard Tobit model. This observation suggests a simple test of exogeneity; using the standard Tobit variance matrix, test to see whether the ρ_k parameters are jointly significantly different from zero. If they are not, then we cannot reject the hypothesis that all the variables are exogenous to the households' choice problem.

The results of Smith and Blundell's (1986) test of exogeneity are provided in the first row of Table 7. For both road and rail demand functions, we can clearly reject the hypothesis that prices and expenditure are exogenous to the choice problem. It appears that we are justified in using an estimation procedure that accounts for endogeneity.

Table 7: Tests for the exogeneity of prices and expenditure in the demand equations

Estimation Method	Test	Test Stat (asymptotic distribution)		p-values	
		Road	**Rail**	**Road**	**Rail**
Control Function	Wald test: $\rho_k = 0$	37.45 $\chi^2(3)$	11.10 $\chi^2(3)$	< .000	.011
Amemiya Generalised Least Squares	Bootstrap Hausman test: $\hat{\delta}_k^{Tobit} = \hat{\delta}_k^{AGLS}$	45.51 $\chi^2(7)$	22.22 $\chi^2(7)$	< .000	.002

To confirm these results, we carry out a second test of exogeneity based on the familiar Hausman test for equivalence of parameters in alternative estimators. Denote by $\hat{\delta}_k^{Tobit}$ the vector of parameter estimates from the Tobit estimator (see Table 4), which are fully efficient if all the explanatory variables are exogenous but inconsistent otherwise. Likewise denote by $\hat{\delta}_k^{AGLS}$ the vector of parameter estimates from the AGLS estimator (see Table 5), which are consistent whether or not the explanatory variables are exogenous. Clearly, if the explanatory variables are exogenous then both the Tobit and AGLS estimators will return consistent estimates that we would expect to be similar to one another. However, if the parameter estimates from the two models differ significantly then we can reject the null hypothesis of exogenous explanatory variables.

The Hausman statistic is based on the vector of contrasts between the two sets of parameters;

$$\tau_k = \left(\hat{\delta}_k^{AGLS} - \hat{\delta}_k^{Tobit}\right) Var\left(\hat{\delta}_k^{AGLS} - \hat{\delta}_k^{Tobit}\right)^{-1} \left(\hat{\delta}_k^{AGLS} - \hat{\delta}_k^{Tobit}\right) \qquad (k = 1, 2) \tag{23}$$

Whilst Hausman (1978) shows how the middle term on the right hand side of (23) (the variance matrix of the vector of differences between the parameters of the two estimators) might be estimated, we choose to employ an alternative procedure. In particular, we apply a bootstrap procedure to estimate $Var\left(\hat{\delta}_k^{AGLS} - \hat{\delta}_k^{Tobit}\right)$. We sample with replacement from the data and re-estimate the Tobit and AGLS models. For each bootstrap sample we calculate the difference between the two vectors of parameters, $\hat{\delta}_k^{AGLS}$ and $\hat{\delta}_k^{Tobit}$. The desired variance matrix is estimated by calculating the empirical variance matrix of the differences resulting from 1,000 replications of the bootstrap procedure.

Using this estimated variance we calculate the Hausman statistic, τ_k $(k = 1, 2)$, which is χ^2 distributed with degrees of freedom equal to the number of parameters in the model. The results for this test are reported in the second row of Table 7. The Hausman test confirms the results of the Smith and Blundell test, rejecting the null hypothesis of exogeneity in the explanatory variables.

Having established that an instrumental variables procedure such as AGLS or the control function approach is required to consistently estimate the parameters, we are interested to test the extent to which the estimated demand relationships for road and rail noise differ. This is not an inconsiderable task. In particular, such a test must account for the fact that the rail and road demand equations are estimated from the same data. As such, any comparison of the two models must account for the fact that the data is not independent; each household contributes an observation to both the rail and road demand data. [57]

To provide a framework within which cross-equation restrictions can be tested, we use an approach adopted by Weesie (2000). First the parameters of the two demand equations are estimated separately using the control function approach. Then the two sets of parameter estimates and associated variance matrices are combined into a single parameter vector and simultaneous variance matrix of the Hubert-White sandwich type that accounts for the multiple observations per household. Using the combined parameter vector and variance matrix it is possible to test cross equation restrictions. The results of a series of such tests are provided in Table 8.

The first seven rows of Table 8 report tests of single parameter constraints, enforcing equality of equivalent parameters across the two demand relationships. The tests reveal that whilst the own price coefficients are statistically indistinguishable, there are significant differences in the coefficients for the cross-price terms, the expenditure term, the constant term and the age factor.

We wish to test to see the extent to which the two demand relationships can be considered as separate estimates of the same demand function for "peace and quiet" from any source. First, we test whether the constant, price and expenditure terms can simultaneously be constrained to take the same values across the two demand relationships. The result is unequivocal; the parameters of the two models are statistically different. Likewise we test to see whether the parameters of the demand shifters can simultaneously be constrained to take the same values across the two demand relationships. Again we can reject this hypothesis with a high level of confidence. Finally, we test the hypothesis that the two demand relationships are identical. Not surprisingly we find that the data does not support this hypothesis.

We conclude that the parameter estimates from our model trace out two statistically distinguishable demand relationships.

---

[57] More efficient estimates of the parameters of the two models could be achieved by exploiting this fact. However, we ignore such complications choosing to follow the recommendation of Amemiya (1979) who suggests that "the gain in asymptotic efficiency does not seem to be worth the added cost of computation" (p.176).

Table 8: Tests of cross equation restrictions on the demand equations

Cross-Equation Parameter Restriction	Test Stat (asymptotic distribution)	p-value
Own price coefficients: $\beta_1^{r_1} = \beta_2^{r_2}$	0.03 $\chi^2(1)$	0.859
Cross price coefficients: $\beta_1^{r_2} = \beta_2^{r_1}$	6.07** $\chi^2(1)$	0.014
Expenditure coefficients: $\beta_1^{\ln(\tilde{m})} = \beta_2^{\ln(\tilde{m})}$	5.34** $\chi^2(1)$	0.021
Constants: $\gamma_1^{cons} = \gamma_2^{cons}$	3.13* $\chi^2(1)$	0.077
Ethnicity factor: $\gamma_1^{ethnic} = \gamma_2^{ethnic}$	1.75 $\chi^2(1)$	0.186
Age factor: $\gamma_1^{age} = \gamma_2^{age}$	24.38*** $\chi^2(1)$	0.000
Family factor: $\gamma_1^{family} = \gamma_2^{family}$	0.93* $\chi^2(1)$	0.335
Constant, Price and Expenditure coefficients: $\gamma_1^{cons} = \gamma_2^{cons}$ & $\beta_1^{r_1} = \beta_2^{r_2}$ & $\beta_1^{r_2} = \beta_2^{r_1}$ & $\beta_1^{\ln(\tilde{m})} = \beta_2^{\ln(\tilde{m})}$	220.88*** $\chi^2(4)$	0.000
Demand shifters: $\gamma_1^{ethnic} = \gamma_2^{ethnic}$ & $\gamma_1^{age} = \gamma_2^{age}$ & $\gamma_1^{family} = \gamma_2^{family}$	35.03*** $\chi^2(1)$	0.000
All coefficients: $\gamma_1^{cons} = \gamma_2^{cons}$ & $\beta_1^{r_1} = \beta_2^{r_2}$ & $\beta_1^{r_2} = \beta_2^{r_1}$ & $\beta_1^{\ln(\tilde{m})} = \beta_2^{\ln(\tilde{m})}$ & $\gamma_1^{ethnic} = \gamma_2^{ethnic}$ & $\gamma_1^{age} = \gamma_2^{age}$ & $\gamma_1^{family} = \gamma_2^{family}$	271.30*** $\chi^2(4)$	0.000

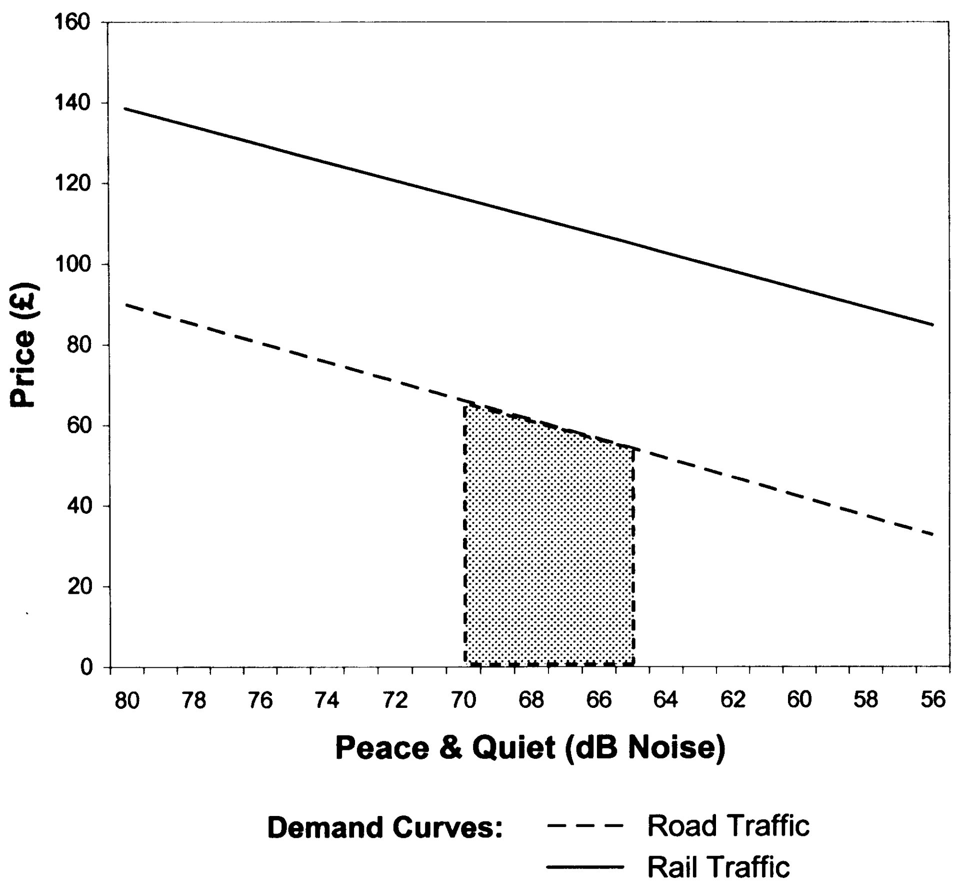
*** Significant at 99% level of confidence
** Significant at 95% level of confidence
* Significant at 90% level of confidence

4. Welfare Estimates

4.i. Introduction

Using the linear demand functions estimated by efficient AGLS, Figure 1 traces the mythical demand curves implied by the coefficient estimates at the means of the covariate data.

Figure 1: Estimated demand curves for Peace and Quiet from Road and Rail Noise Pollution at the means of the covariate data



Whilst the slopes on the two linear demand curves are very similar, the demand curve for peace and quiet from rail traffic noise is considerably higher than that from road traffic noise. The graphs illustrates the conclusions of the formal tests carried out in the previous section; despite the use of a common unit of measure dB_{LEQ}, households do not consider these two sources of noise to be equivalent. Indeed, the demand equations estimated here indicate that household's are willing to pay more to avoid rail traffic noise than they are to avoid road traffic noise. One possible explanation of such an observation is that the common unit used to measure noise from these two sources (dB_{LEQ}) conceals important differences in the characteristics of road and rail noise.

4.ii. Welfare Estimates

The benefits of an external change in the provision of a good can be approximated by the area under the demand curve between the old level of provision and the new level of provision. As described in the last chapter, this area should return a value that does not differ greatly from the quantity compensating zurplus (QCS) measure of welfare. For example, the shaded trapezium in Figure 1 defines the welfare benefits resulting from a reduction in road traffic noise pollution from 70dB to 65dB.

The actual value of this welfare gain can be read off from Table 9. This table provides welfare values for 1dB changes in noise pollution levels from road traffic and rail traffic (columns 3 and 5 respectively) calculated at the mean values of the covariate data. To calculate the welfare benefits from larger changes simply requires the relevant range of 1dB changes to be summed. This is illustrated in columns 4 and 6 of Table 9 where the benefits of particular 5dB changes are listed. The welfare estimates are values per annum reported in 1997 prices.

Using this information we can put a figure to the shaded area in Figure 1; the welfare benefits of a reduction in road traffic noise pollution from 70dB to 65bB is estimated to be around £300 per year.

The values for road noise range from £31.49 per annum for a 1dB reduction from a 56dB baseline to £88.76 per annum for the same change from a 80dB baseline. The equivalent values for rail noise are somewhat higher ranging from £83.61 to £137.41 per annum. As far as the author is aware, these are the first published estimates of welfare values for noise pollution changes derived from hedonic property market data in a theoretically correct manner. As such there is no comparative literature with which to assess their reliability. Suffice it to say that the magnitude and range of values listed in Table 9 appear entirely plausible to the author.

Table 9: Welfare estimates for changes in noise exposure at means of covariate data (95% confidence interval range)

Noise Change		Welfare Change per Annum (1997 prices)			
High	**Low**	**Road**		**Rail**	
56	55	31.49 (24.84 to 52.52)		83.61 (43.21 to 461.80)	
57	56	33.88 (26.51 to 57.26)		85.85 (44.43 to 473.39)	
58	57	36.26 (28.18 to 61.97)	181.32	88.09 (45.65 to 484.98)	440.45
59	58	38.65 (29.81 to 66.64)		90.33 (46.87 to 496.58)	
60	59	41.04 (31.48 to 71.31)	(140.90 to 309.87)	92.57 (48.08 to 508.17)	(228.24 to 2,424.92)
61	60	43.42 (33.14 to 75.97)		94.82 (49.30 to 519.77)	
62	61	45.81 (34.81 to 80.64)		97.06 (50.52 to 531.36)	
63	62	48.19 (36.48 to 85.31)	240.97	99.3 (51.73 to 542.96)	496.49
64	63	50.58 (38.15 to 89.97)		101.54 (52.95 to 554.55)	
65	64	52.97 (39.82 to 94.77)	(182.41 to 426.53)	103.78 (54.17 to 566.14)	(258.67 to 2,714.78)
66	65	55.35 (41.49 to 99.57)		106.02 (55.39 to 580.05)	
67	66	57.74 (43.16 to 104.37)		108.26 (56.60 to 594.85)	
68	67	60.12 (44.82 to 109.17)	300.62	110.51 (57.82 to 609.65)	552.53
69	68	62.51 (46.50 to 113.97)		112.75 (59.04 to 624.45)	
70	69	64.9 (48.16 to 118.77)	(224.12 to 545.85)	114.99 (60.26 to 639.25)	(289.10 to 3,048.25)
71	70	67.28 (49.83 to 123.57)		117.23 (61.47 to 654.04)	
72	71	69.67 (51.50 to 128.37)		119.47 (62.69 to 668.84)	
73	72	72.05 (53.17 to 133.18)	360.27	121.71 (63.91 to 683.64)	608.57
74	73	74.44 (54.84 to 137.98)		123.96 (65.12 to 698.44)	
75	74	76.83 (56.50 to 142.78)	(265.84 to 665.88)	126.2 (66.34 to 713.24)	(319.53 to 3,418.20)

76	75	79.21 (58.17 to 147.58)		128.44 (67.47 to 728.03)	
77	76	81.6 (59.84 to 152.38)		130.68 (68.52 to 742.83)	
78	77	83.98 (61.51 to 157.18)	419.92	132.92 (69.57 to 757.63)	664.61
79	78	86.37 (63.18 to 161.98)		135.16 (70.62 to 772.43)	
80	79	88.76 (64.85 to 166.78)	(307.55 to 785.90)	137.41 (71.67 to 787.23)	(347.83 to 3,788.16)

Furthermore, Table 9 reports 95% confidence intervals for each of the reported welfare estimates. These are calculated using a non-parametric bootstrap procedure. We sample with replacement from the data and re-estimate both demand relationships using the bootstrap sample. For each bootstrap sample we calculate an estimate of the welfare value for each of the noise changes listed in Table 9. The confidence intervals reported in Table 9 are the 2.5th and 97.5th percentile values from the distribution of welfare estimates resulting from 1,000 replications of the bootstrap procedure.

It is immediately apparent that the 95% confidence intervals for the rail welfare estimates are considerably wider than those for the road welfare estimates. Of course, this is only to be expected given the fact that our dataset contains many fewer observations of exposure to rail traffic noise above the urban background level than is the case for road traffic noise.

Moreover, notice that across the range of noise changes reported in Table 9, there is considerable overlap in the confidence intervals for the road and rail welfare values. Furthermore, for small changes at low levels of noise exposure, say a reduction from 56 to 55 dB, the confidence intervals overlap much less than they do for small changes at high levels of noise exposure, say a reduction from 76 to 75 dB. This pattern suggests that our estimates of the intercept parameters of the demand relationships are more precise than our estimates of the slopes of the demand curves. Again, this is to be expected. The intercept term is largely determined by observations of households choosing properties in relatively quiet locations. Conversely, the slope of the demand function is largely determined by observations of households choosing properties in relatively noisy locations. Since the dataset is rich in observations at low levels of noise and relatively sparse in observations at high levels of noise, it is not surprising to observe that the intercept is estimated with relatively more precision than the slope parameter.

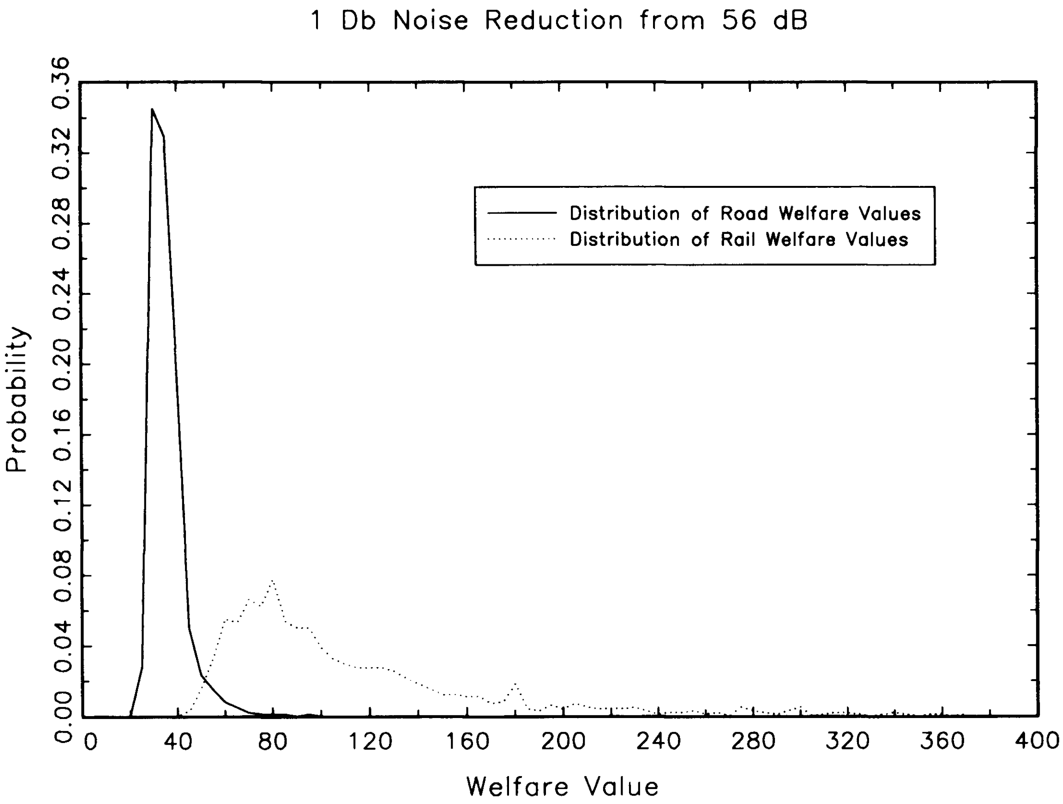
To better appreciate the precision with which the welfare values have been estimated observe Figure 2. This presents plots of the distributions of the 1,000 bootstrapped welfare values estimated for 1 dB reductions in rail and road noise pollution from two different baseline levels of noise exposure. To be clear, these figures give an indication of the precision with which the mean welfare values reported in Table 9 are estimated. In particular, they provide a nonparametric estimate of the distribution of the mean welfare estimates for the entire Birmingham sample and not the distribution of welfare values across different households in the sample.

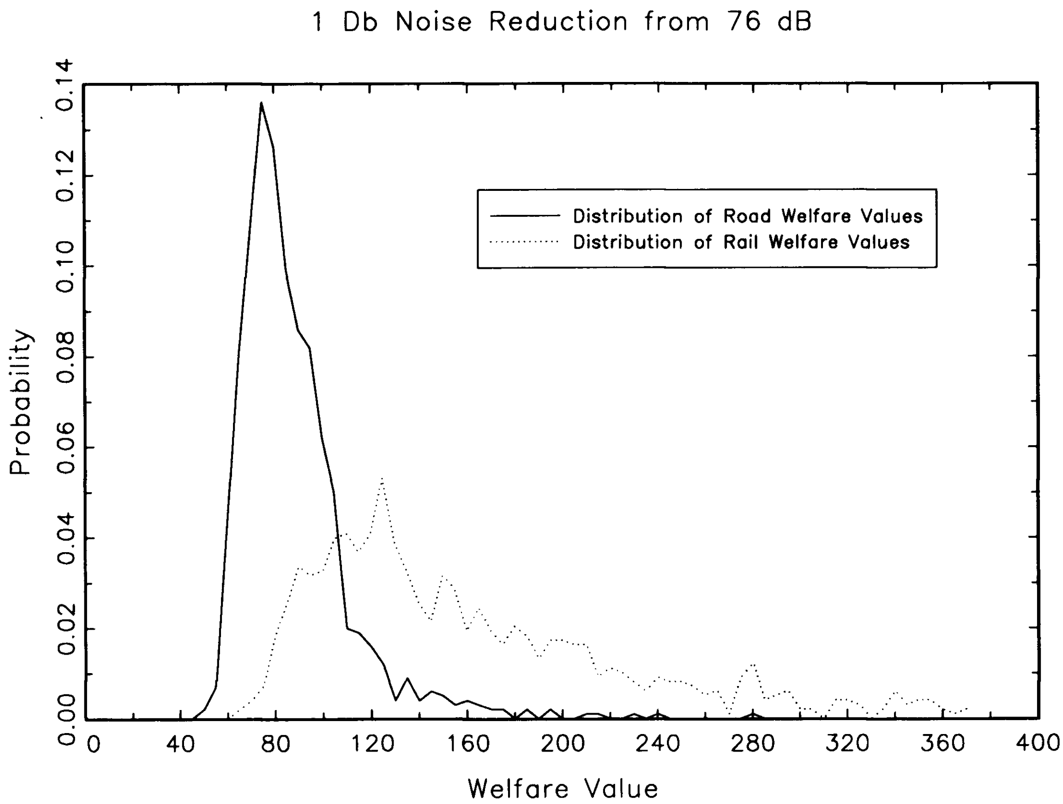
The patterns discussed with reference to the 95% confidence intervals are immediately apparent from the distributions plotted in Figure 2. Notice that the distribution of the mean welfare estimate for changes in road noise is considerably more concentrated than that for rail noise, reflecting the greater precision with which the road noise demand equation has been estimated. Also the variance of the mean welfare estimates for both road and rail noise are smaller for a 1dB change from a 56dB baseline than they are for a 1dB change from a 76dB baseline. Notice also that the distribution of the mean welfare estimates for road and rail noise changes overlap and that this overlap is greater at higher baselines than it is at lower baselines.

Of course, if the 95% confidence intervals overlap, then it might be the case that households enjoy the same welfare benefit from a change in their exposure to road traffic noise as they do to a change in their exposure to rail traffic noise. Column 3 of Table 10 provides a test of this hypothesis. Using evidence from the bootstrap, we estimate the probability that the road and rail welfare estimates are the same. For small changes below a 65dB baseline we can reject the hypothesis that the rail and road welfare estimates are the same with at least 95% confidence. For changes from higher baselines we cannot be so certain that the welfare estimates are different.

Given the possibility that households enjoy the same welfare benefit from a change in road noise as they do from the same change in rail noise, we develop a procedure to suggest a best "combined" value applicable to both road and rail noise. This best "combined" figure is reported in Column 4 of Table 10.

Figure 2: Bootstrap distribution of mean road and rail welfare estimates for two different changes in the noise environment





The combined figures are estimated from the bootstrap distributions such as those in Figure 2. In particular, we grid-search across welfare values and for each value calculate the probability mass falling into the upper tail of the road noise distribution above this welfare value and the probability mass falling into the lower tail of the rail noise distribution below this welfare value. These two probabilities (reported in the final two columns of Table 10) indicate the likelihood that the road noise welfare estimate is at least as high, and that the rail noise welfare estimate is at least as low as the test value. We select the combined value as that which maximises the product of these two likelihoods.

Table 10: Results of bootstrap comparison of welfare estimates for rail and road noise exposure at means of covariate data

Noise Change		Prob. Road & Rail Welfare Values differ	Best Estimate of "Combined" Road & Rail Value		
High	Low		"Combined" Welfare Value	Prob. Road >= "Combined"	Prob. Rail <= "Combined"
56	55	0.026	62.67	0.010	0.134
57	56	0.028	61.40	0.018	0.105
58	57	0.030	59.94	0.034	0.076
59	58	0.032	65.34	0.032	0.111
60	59	0.034	68.96	0.034	0.131
61	60	0.036	73.49	0.034	0.155
62	61	0.038	78.04	0.034	0.183
63	62	0.044	82.54	0.034	0.215
64	63	0.047	86.96	0.034	0.242
65	64	0.050	85.76	0.045	0.207
66	65	0.056	80.11	0.075	0.140
67	66	0.064	80.41	0.091	0.128
68	67	0.071	74.13	0.195	0.069
69	68	0.074	76.92	0.201	0.077
70	69	0.079	80.28	0.195	0.088
71	70	0.084	82.93	0.201	0.096
72	71	0.090	86.42	0.195	0.111
73	72	0.096	88.70	0.202	0.117
74	73	0.098	92.46	0.196	0.132
75	74	0.102	93.92	0.210	0.130
76	75	0.106	98.60	0.196	0.150
77	76	0.113	95.06	0.278	0.112
78	77	0.122	98.92	0.260	0.131
79	78	0.131	100.51	0.282	0.130
80	79	0.140	102.17	0.296	0.130

4.iii. Benefits transfer

Since the values in Table 9 are derived from the underlying preferences defined by a demand curve there is no theoretical impediment to their use in benefits transfer exercises. Unfortunately, two issues complicate the simple use of the reported values;

- For changes in the noise environment starting from a high baseline, we find that we cannot reject the hypothesis that road traffic and rail traffic welfare estimates are the same.
- The values in Table 9 are those enjoyed by a household exhibiting the mean characteristics of households in the Birmingham dataset. Clearly, we might wish to adjust our welfare estimates to reflect the differing socioeconomic characteristics of households at the study site.

With regards to the former point, our testing of the demand relationships suggests that there are statistically significant differences in households' preferences for avoidance of these two sources of noise pollution. However, these differences are not estimated with sufficient precision at high baseline levels of noise to ensure that the welfare estimates themselves are always significantly different. From a statistical point of view, the results posted in Table 9 represent our best estimates of the true welfare benefits of changing noise environments, and it would be recommended that these be used in policy decisions. However, if a "combined" welfare estimate that does not distinguish between the two sources of noise were preferred for policy purposes, such a figure is provided in Table 10.

With regards to the latter point, assuming that the population at the transfer site has identical characteristics to those exhibited at the mean of the Birmingham data is an assumption we may wish to relax. Certainly we may wish to adjust our welfare estimates to reflect the role of income in determining WTP.

Tables 11 and 12 list welfare estimates at the 5^{th}, 25^{th}, 50^{th}, 75^{th} and 95^{th} percentiles of the property expenditure distribution for the Birmingham data. For both road and rail traffic noise, welfare values increase by approximately £50 per annum moving from low-income households sitting at the 5^{th} percentile of the expenditure distribution to high-income households sitting at the 95^{th} percentile. This relatively large value reinforces our assertion that accounting for variation in income/wealth is an important practice in benefits transfer exercises.

A number of difficulties with the application of these results in benefit transfer exercises suggest themselves. First, the expenditure variables used in the analysis are mythical not actual expenditures. In performing a benefits transfer, we will not know the mythical expenditures of households in the transfer location. Fortunately, the differences between actual and mythical expenditure turn out to be relatively minor. We find that 90% of households in the sample have their expenditures adjusted by less than 3.1% and the maximum adjustment is some 25.3%. As such, we contend that using the actual property expenditures of households at the transfer location will not result in undue errors in the calculation of welfare estimates.

The second difficulty in applying these results is the existence of property price inflation. Property prices in Birmingham in 2004 are two to three times what they were in 1997. Unfortunately, under an assumption of two-stage budgeting (see Section 2 of this Chapter) it is this property price (total property expenditure) that enters the demand equation. As such, if one were to take the estimated demand functions at face value, then recalculation of welfare benefits for 2004 would require the property price argument to be increased two to three times. Inspection of the Birmingham data set reveals that this is equivalent to a move from the 10^{th} percentile to the 90^{th} percentile of the distribution of property prices in 1997 In other words, welfare values should now be some £20 per dB per annum greater for road noise and some £50 per dB per annum greater for rail noise than they were back in 1997. Clearly, this is unrealistic. Indeed, apart from the fact that the values in Tables 11 and 12 are in 1997 prices, we see no reason why households' valuations of 'peace and quiet' should have increased by anything like this much between 1997 and the present.

As such we recommend a pragmatic alternative. First we note that a household's expenditure on property is very closely related to their level of income. As such, those at the 95^{th} percentile of the Birmingham property expenditure distribution are also likely to be near the 95^{th} percentile of the Birmingham income distribution. Our suggestion is to match the transfer site with the welfare values in Tables 11 and 12 using measures of income rather than measures of property expenditure. That is, transfer exercises should attempt to match the income of households at the transfer site with the equivalent percentile of the income distribution in the city of Birmingham. For example, a location characterised by household incomes that fell on the 75th percentile of the Birmingham income distribution should use the welfare values in the 75th percentile columns of Tables 11 and 12.

One source of error in such a procedure is that our data relate only to households that purchase properties. As such, the assumption that the distribution of property expenditures in Birmingham can be taken as reflecting the distribution of incomes in Birmingham should really be qualified. That is to say, the distribution of expenditures can only be thought to reflect the distribution of incomes of *property-purchasing* households in Birmingham.

An even simpler alternative would be to assume that the income distribution in Birmingham is reasonably typical of the UK as a whole, in which case benefits transfer exercises need only determine the position of the transfer population on the UK income distribution.

The welfare values described in Tables 11 and 12 are quoted in 1997 prices. When transferring these values across time, the estimates should be updated to account for price inflation.

Appendix G contains tables equivalent to 11 and 12 but for percentiles of the socioeconomic characteristic distributions. Here we are less eager to recommend the application of adjustments to welfare estimates in benefits transfer procedures. This reluctance stems from three sources.

First, calculation of factor scores for a transfer site is a somewhat involved procedure; data on a number of variables have to be collated from the census, these must be manipulated to create the variables used in the factor analysis, then combined in a linear combination using the coefficients listed in Table B1 of Appendix B. Whilst this procedure is by no means impossible, there are other reasons why we think this unnecessary. Prime amongst these is that the role of the factors in the demand equations is to act as crude proxies for missing data on household characteristics. As such the coefficient estimates on these factors are somewhat difficult to interpret. Consequently it would be difficult to find an intuitive defence of benefits adjustments with regard to factor scores in the same way one can convincingly argue that those on higher income are willing to pay more for reductions in noise pollution. Finally and reassuringly, differences in the factor scores have, in the main, only a limited impact on welfare values. As can be seen from the tables in Appendix G, moving from the 5th to the 95th percentile of the ethnicity factor distribution adds only £12 to the road traffic pollution values and only £3 to the rail traffic pollution values. Similarly, moving from the 5th to the 95th percentile of the family composition factor distribution only adds around £7.50 to welfare values from changes in both road and rail traffic noise. Moreover, moving from the 5th to the 95th percentile of the age composition factor increases welfare values from road noise reductions by only £7. The only exception to the relative insignificance of the factor scores in determining welfare values is for family composition in the rail noise demand equation. Here moving from the 5th to 95th percentile precipitates a £24 change in welfare values. However, even this relatively large difference in values is only half of that occasioned by a comparison of welfare values at the 5th to 95th percentiles of the expenditure distribution.

In conclusion the author recommends that Tables 11 and 12 be used for benefits transfer exercises. To simplify such transfers it is suggested that investigators determine the percentile of the UK income distribution best describing the population at the transfer location, then adopt the figures in the appropriate percentile column of Tables 11 and 12 as best estimates of the welfare values at the transfer site.

Table 11: Welfare estimates for changes in Road noise exposure at different percentiles of the Expenditure Distribution

Noise Change (dB)	Welfare Values for Changes in Road Noise at Percentiles of Expenditure Distribution (£ per annum in 1997 prices)						
	5th	10th	25th	50th	75th	90th	95th
56 to 55	18.72	21.84	26.11	30.64	36.05	42.54	47.50
57 to 56	21.11	24.23	28.49	33.03	38.43	44.92	49.89
58 to 57	23.50	26.61	30.88	35.41	40.82	47.31	52.27
59 to 58	25.88	29.00	33.27	37.80	43.20	49.70	54.66
60 to 59	28.27	31.39	35.65	40.19	45.59	52.08	57.05
61 to 60	30.66	33.77	38.04	42.57	47.98	54.47	59.43
62 to 61	33.04	36.16	40.42	44.96	50.36	56.85	61.82
63 to 62	35.43	38.54	42.81	47.34	52.75	59.24	64.20
64 to 63	37.81	40.93	45.20	49.73	55.13	61.63	66.59
65 to 64	40.20	43.32	47.58	52.12	57.52	64.01	68.98
66 to 65	42.59	45.70	49.97	54.50	59.91	66.40	71.36
67 to 66	44.97	48.09	52.35	56.89	62.29	68.78	73.75
68 to 67	47.36	50.47	54.74	59.27	64.68	71.17	76.14
69 to 68	49.74	52.86	57.13	61.66	67.06	73.56	78.52
70 to 69	52.13	55.25	59.51	64.05	69.45	75.94	80.91
71 to 70	54.52	57.63	61.90	66.43	71.84	78.33	83.29
72 to 71	56.90	60.02	64.28	68.82	74.22	80.71	85.68
73 to 72	59.29	62.40	66.67	71.20	76.61	83.10	88.07
74 to 73	61.67	64.79	69.06	73.59	78.99	85.49	90.45
75 to 74	64.06	67.18	71.44	75.98	81.38	87.87	92.84
76 to 75	66.45	69.56	73.83	78.36	83.77	90.26	95.22
77 to 76	68.83	71.95	76.21	80.75	86.15	92.64	97.61
78 to 77	71.22	74.33	78.60	83.13	88.54	95.03	100.00
79 to 78	73.60	76.72	80.99	85.52	90.92	97.42	102.38
80 to 79	75.99	79.11	83.37	87.91	93.31	99.80	104.77

Table 12: Welfare estimates for changes in Rail noise exposure at different percentiles of the Expenditure Distribution

Noise Change (dB)	Welfare Values for Changes in Rail Noise at Percentiles of Expenditure Distribution (£ per annum in 1997 prices)						
	5th	10th	25th	50th	75th	90th	95th
56 to 55	51.78	59.55	70.19	81.49	94.96	111.15	123.52
57 to 56	54.02	61.79	72.43	83.73	97.20	113.39	125.77
58 to 57	56.26	64.03	74.67	85.97	99.44	115.63	128.01
59 to 58	58.50	66.28	76.91	88.21	101.69	117.87	130.25
60 to 59	60.75	68.52	79.15	90.46	103.93	120.11	132.49
61 to 60	62.99	70.76	81.39	92.70	106.17	122.35	134.73
62 to 61	65.23	73.00	83.63	94.94	108.41	124.60	136.97
63 to 62	67.47	75.24	85.88	97.18	110.65	126.84	139.22
64 to 63	69.71	77.48	88.12	99.42	112.89	129.08	141.46
65 to 64	71.95	79.72	90.36	101.66	115.13	131.32	143.70
66 to 65	74.20	81.97	92.60	103.91	117.38	133.56	145.94
67 to 66	76.44	84.21	94.84	106.15	119.62	135.80	148.18
68 to 67	78.68	86.45	97.08	108.39	121.86	138.05	150.42
69 to 68	80.92	88.69	99.33	110.63	124.10	140.29	152.67
70 to 69	83.16	90.93	101.57	112.87	126.34	142.53	154.91
71 to 70	85.40	93.17	103.81	115.11	128.58	144.77	157.15
72 to 71	87.65	95.42	106.05	117.36	130.83	147.01	159.39
73 to 72	89.89	97.66	108.29	119.60	133.07	149.25	161.63
74 to 73	92.13	99.90	110.53	121.84	135.31	151.50	163.87
75 to 74	94.37	102.14	112.78	124.08	137.55	153.74	166.12
76 to 75	96.61	104.38	115.02	126.32	139.79	155.98	168.36
77 to 76	98.85	106.62	117.26	128.56	142.03	158.22	170.60
78 to 77	101.09	108.87	119.50	130.81	144.28	160.46	172.84
79 to 78	103.34	111.11	121.74	133.05	146.52	162.70	175.08
80 to 79	105.58	113.35	123.98	135.29	148.76	164.94	177.32

5. Conclusions

This final chapter reports the policy relevant research findings of this thesis. These are contained in Tables 11 and 12 which report welfare values for changes in noise pollution from road and rail traffic respectively. The figures in these tables are derived from the estimation of demand equations for peace and quiet. As far as the author is aware, these are the first welfare estimates for noise pollution to be derived from a hedonic analysis in a theoretically consistent manner.

A number of data and econometric issues have been overcome in the estimation of these equations, most notably in handling large scale censoring of the dependent variable and the endogeneity of prices and expenditure.

Whilst there is little published material with which to compare these estimates, the values appear to be of very plausible magnitudes. For road noise above the assumed urban background level of 55dB, these range from a low of £19 per annum (in 1997 prices) for a low-income household experiencing a 1dB change in a relatively quiet environment (56dB), to a high of £105 per annum for a high-income household experiencing a 1dB change in a noisy environment (80dB). For rail noise above the background level, the values are somewhat higher ranging from £52 per annum for a low-income household experiencing a 1dB change in a relatively quiet environment (56dB) to £178 per annum for a high-income household experiencing a 1dB change in a noisy environment (80dB).

It is argued that the figures in Tables 11 and 12 be used in benefits transfers exercises by equating the percentiles of the distribution of property expenditures listed there with the percentiles of the UK income distribution.

PART 4

CONCLUSIONS, ACHIEVEMENTS AND FUTURE DIRECTIONS

CHAPTER 10: CONCLUSIONS, ACHIEVEMENTS AND FUTURE DIRECTIONS

1. Introduction

This thesis describes the application of hedonic pricing techniques to data collected from the City of Birmingham property market. Put simply, the principal objective of the research has been to examine differences in property prices in Birmingham to determine how much households are willing to pay to avoid changes in their exposure to noise pollution.

In fulfilling that principal objective the author has sought to achieve a number of goals. In particular, the author has attempted to firmly root the research presented in this thesis in a solid foundation of *economic theory*. It is hoped that Parts 1 and 3 of this thesis contain an articulate account of that theory, clearly explicating important theoretical considerations that have been explained poorly elsewhere, providing some small additions to that theoretical literature and summarising some of the very latest developments in the theory of hedonic pricing. Moreover, the particular strengths of the author lie in the field of *applied econometrics*. As such, the empirical work presented in this thesis has endeavoured to achieve the highest standards in that field. The author believes that the most important contributions of this thesis lie in the pioneering of new techniques of data analysis described in Part 2. Finally, the author has sought to produce a piece of work with *policy-relevant outputs*. The estimates of welfare values relating to changes in exposure to traffic-related noise pollution reported in the final Chapter of Part 3 of this thesis, represent the product of that endeavour.

The purpose of this final Chapter is to assess the extent to which the author has been successful in those aims. Under the three headings of economic theory, applied econometrics and policy-relevant outputs, this Chapter considers the strengths and weaknesses of the work presented in this thesis and describes some directions that future research in this field might take.

2. Theory of Hedonic Pricing

2.i. Models of Equilibrium in Hedonic Markets

Chapter 1 of this thesis sets out in some detail the theoretical model of market equilibrium in property markets that underpins the method of hedonic pricing. Whilst offering no new theoretical developments, Chapter 1 attempts to provide a clear and accessible presentation of this occasionally confusing literature. More importantly, the exposition includes discussion of the very latest advances in the theoretical modelling of hedonic markets. For want of space, the details of these models have been relegated to Appendix A.

A key focus of these recent advances has been to examine the nature of the equilibrium property market and its associated HPF when simplifying assumptions concerning the distribution of household preferences and the supply of properties to the market are relaxed. One of the insights afforded by the work presented in this thesis is an examination of the implications of these theoretical advances for applied work in the field of hedonic pricing. In particular, this examination identifies three predictions of the theoretical models that have implications for applied work;

- *Nonlinearity of the HPF.* The models predict that the equilibrium hedonic price schedule will likely be a highly nonlinear function of property attributes.
- *"Lumpy" provision in attribute space.* The theoretical models observe that in equilibrium, the market may not provide a continuum of products over attribute space. Depending on the model, this phenomenon manifests itself in one of two ways. In one set of models it is found that the equilibrium market will contain clusters of properties exhibiting similar combinations of attributes, whilst properties with other combinations of attributes are relatively sparse. In another set of models it is found that, in equilibrium, neighbourhoods with certain socioeconomic mixes are common whilst other combinations are relatively rare.

Whilst numerous studies have considered nonlinearity of the HPF, few if any have explored the implications of sorting and "lumpy" provision. One original proposal investigated in this thesis is that these latter characteristics of property markets suggest a procedure for estimating highly nonlinear HPFs from empirical data. In particular an estimation strategy is proposed in which data-driven clustering techniques are employed to identify concentrations of properties displaying similar attributes. Since properties in each of these concentrations lie in close proximity in attribute space, they must, by definition, lie close to each other on the hedonic price surface. The proposed estimation strategy draws on this observation. Whilst the standard approach has been to employ increasingly general econometric specifications to capture the nonlinearity of the equilibrium HPF over all attribute space, the estimation strategy proposed here is to fit separate price functions for the properties in each cluster. Similar to spline fitting, each estimated price function forms a local approximation to the hedonic price surface over the attribute area spanned by the properties in each cluster.

Chapter 6 provides an application of this proposed estimation approach. Interestingly, it is found that identifying clusters according to the socioeconomic characteristics of neighbourhoods leads to a better fitting approximation to the HPF than identifying clusters according to the structural attributes of properties. It seems that the HPF possesses greater nonlinearity in dimensions reflecting characteristics of a properties' neighbourhood than in dimensions indicating structural characteristics.

As described in Chapters 3 and 6 there is a long tradition in hedonic analysis of partitioning property market data and estimating separate HPFs for each partition. Previously, partitioning has been justified through the assumption of market segmentation. Each partition of the data is assumed to represent a separate market segment characterised by a separate HPF. Rarely, however, do authors elucidate on the market barriers or processes that precipitate such segmentation. The advance in thinking provided in this thesis is that partitioning the data is motivated by the predictions of theoretical models indicating that property markets in equilibrium will be characterised by lumpy provision in attribute space. The existence of such clusters of properties is distinct from the notion of market segmentation. As such, the approach used to identify clusters need not be shackled by the requirement of providing a formal definition of the process driving market segmentation or to formally define the property or neighbourhood characteristics by which such segments should be delineated. Rather in the work presented here, the data itself is used to inform on the pattern of clustering in the property market.

The application reported in Chapter 6 represents an initial attempt to put the partitioning strategy into practice. The limitations of that application lie in the fact that the clusters of properties for which HPFs are estimated, are defined only on subsets of the attribute space. That is, clusters are defined either by neighbourhood characteristics or by structural characteristics. As such the estimated models only allow for nonlinearities in one particular set of dimensions of attribute space. Moreover, the simple functional forms used to approximate the HPF over each cluster's attribute space do not allow for much flexibility in the HPF over that restricted space.

All the same, the author sees much potential in the partitioning estimation approach. In particular, future research would seek to exploit the obvious similarities between the proposed approach and spline fitting techniques (Fuller, 1969; Smith, 1974). The latter technique represents the fit as a spline; a piecewise polynomial. The regions that define the pieces are separated by a series of knots or breakpoints. Clearly, these regions bare close resemblance to the attribute spaces identified as clusters in the analysis in Chapter 6. Indeed the well-developed theory of spline fitting provides many potential avenues of research. In particular, the author would like to investigate the use of multivariate adaptive regression splines (MARS) (Friedman, 1991) a spline fitting technique which seeks to automatically determine which regions of the attribute space require separate splines.

As a final comment, there is a received wisdom in the field of hedonic pricing that theoretical models are of little use in directing empirical research. It is the author's opinion that this is far from the truth. Rather, it is increasingly apparent that theory offers numerous insights that may be of interest to the applied researcher. Indeed, applied researchers might look to the theoreticians to provide them with guidance on a number of issues that remain relatively unstudied. For example, models of equilibrium in property markets generally assume a complete and continuous attribute space. That is, properties displaying any combination of attributes are possible (if not necessarily made available) in the property market. Applied researchers could learn much from theoretical models that examined the nature of the equilibrium market and its associated HPF when the attribute space is bounded or contains holes. Similarly, there appears to have been little theoretical work examining the process of adjustment in the market precipitated by exogenous changes in the provision of property attributes (e.g. the provision of new transport infrastructure or the introduction of policy measures that reduce noise pollution). Again theoretical work examining the dynamics of market adjustment would be of interest to those interested in the distributional impacts of policy.

2.ii. Welfare Estimation in Hedonic Markets

A second area in which this thesis has concerned itself with the theoretical underpinnings of hedonic pricing is in defining and measuring welfare changes precipitated by changes in environmental quality.

Chapter 2 of this thesis provides a thorough and coherent account of the variety of ways in which welfare changes might be measured in property markets. Some time is spent extending descriptions of these measures to contexts where property owners are constrained in their responses to environmental change.

Unfortunately, given the complexity of adjustments in the property market following a change in environmental quality, few of the welfare measures defined in the theoretical literature could realistically be quantified in real world contexts. Nevertheless, the author invokes an argument clearly elucidated by Bartik (1988) to justify focussing attention on estimating demand curves for property attributes such as environmental quality. As is well-known, welfare values are customarily measured as areas under such curves, though the story is not quite so simple in property markets. All the same, Bartik showed that summing up the welfare values calculated in this way for all households directly affected by an environmental improvement provides a lower bound estimate to the total long term welfare change experienced by all economic agents in the property market.

The discussion in Chapter 2 takes time to stress that this measure is only valid in the aggregate, a point frequently overlooked in the hedonic pricing literature. In particular this measure tells us nothing about the long term benefits accruing to specific households once all market adjustments have taken place.

Numerous difficulties complicate the actual estimation of demand functions from real world property market data. For example, data from one market only provides information pertaining to one point on the demand function for a particular household. One possible solution is to estimate HPFs for several different property markets. By observing the choices of similar households facing different pricing schedules in the alternative markets, sufficient information can be collected to approximate the slope of the demand function. Moreover, the process of recovering an estimate of the demand curve is complicated by the fact that in hedonic markets, prices per unit of a characteristic are not necessarily constant. Under such circumstances, households effectively simultaneously choose the price and quantity of the property characteristic such that prices must be treated as endogenous in the estimation process. The author's responses to these various methodological difficulties are discussed later in this Chapter.

The final issue pertaining to economic theory considered in this thesis is that of the relationship between *ordinary, compensated,* and *mythical demand functions*. As is well-known, areas under the compensated demand curve provide exact measures of welfare change. Unfortunately, it is the ordinary demand function, not the compensated demand function, that can be estimated from observed data. All the same, under reasonable assumptions concerning the income elasticity of demand, the ordinary and compensated functions for goods traded in standard markets are sufficiently similar for the differences between the two to be ignored (these issues are discussed by Willig, 1976, Randall and Stoll, 1980, Hanemann, 1991, and Sugden, 1999, amongst others).

Whilst welfare estimation using the ordinary demand curve is a reasonable approximation in standard markets, the same may not be true of hedonic markets. In particular, when marginal prices are not constant the usual manifestation of the ordinary demand function as a downward sloping curve relating quantities chosen to marginal prices does not exist. Accordingly, it is no longer appropriate (nor for that matter possible) to approximate welfare measures by calculating areas under the ordinary demand curve. Whilst these issues have been addressed previously in the literature, Chapter 8 provides a thorough and accessible account of this oft-ignored issue. As the discussion in Chapter 8 highlights, hedonic pricing studies that fail to recognise and account for this fundamental issue are necessarily flawed.

One well-established procedure for dealing with this problem is to make adjustments to the income argument in the demand relationship so as to estimate the *mythical demand function*. This function reveals how the ordinary demand function would look if the prices faced by households were actually constant at the implicit prices selected by those households in purchasing their optimal properties. As such the mythical demand function exhibits the familiar downward sloping relationship between prices and quantities.

The critical question for welfare analysis is the extent to which mythical demand curves approximate compensated demand curves. In this case the relationship between the two not only depends on income effects, but also on the degree of curvature of the HPF. Theoretical work by Edlefsen (1981) and Blomquist (1989) has sought to describe the relationships that exist between the mythical and ordinary demand functions when marginal prices are not constant. Unfortunately, there appears to be no theoretical research comparing the mythical and compensated demand curves in a world of non-constant marginal prices to parallel that comparing the ordinary and compensated demand functions in a world of constant marginal prices. The author acknowledges that given this deficiency in the theoretical literature, relying on mythical demand curves for welfare estimation is a weakness of the current study and highlights this as an area worthy of future research.

In matter of fact, Palmquist (1988) suggests an approach that might be used to overcome these difficulties. Palmquist's approach involves assuming a particular functional form to describe the underlying preferences of the household. Using well-known relationships in demand theory one can use this assumed preference structure to derive an expression for the mythical demand function. Estimating this expression for the mythical demand function using real world data would allow one to recover the parameters of the assumed preference structure and hence the compensated demand curve. The drawback of Palmquist's approach is that by assuming a functional form for preferences, the researcher imposes a great deal of untestable structure on the econometric model. Again, the author acknowledges that Palmquist's approach to identifying the structure of preferences is worthy of further investigation.

3. Applied Econometrics in Hedonic Pricing

The hedonic pricing method progresses in two stages. In the first stage one examines data from a number of property markets to determine the price that must be paid in each of those markets for each extra unit of an environmental good. In this thesis, the environmental good under consideration is 'peace and quiet'. In the second stage data from each of the markets is combined and used to estimate the (mythical) demand function for 'peace and quiet'. In particular, variation in the price of 'peace and quiet' between markets allows one to establish how households' demand for 'peace and quiet' responds to differences in price. The bulk of this thesis, therefore, reports on an exercise in applied econometrics; first in estimating the HPF and second in estimating mythical demand functions. In this section, the author reviews some of the achievements, weaknesses and future research directions pertaining to that exercise in applied econometrics.

3.i. Partitioning property market data

As described in Section 2 of this Chapter, there are two justifications that might be forwarded for partitioning property market data from a single urban area and estimating separate HPFs for each partition. The first justification maintains that the separate partitions represent separate market segments whilst the second justification, outlined in detail above, supports partitioning as a means of approximating a highly nonlinear HPF. Either way, an appropriate method for partitioning the data must be found.

In this study the author has employed a method of analysis called *model-based clustering*, a statistically rich set of analytical tools that use the data itself to inform on the pattern of segmentation in the property market. These techniques have never previously been applied in this context and, as far as the author is aware, have never been applied by econometricians in the analysis of any economics data set.

The great advantage of model-based clustering, over other clustering approaches, is that it uses statistical techniques to determine the number and nature of distinct groups present in the data. Furthermore, it allows for observations to be allocated to a noise category; that is the procedure allows for observations that cannot readily be classified as belonging to any particular cluster.

As always, there are some drawbacks to the model-based clustering approach. In particular the algorithm through which the clustering problem is solved requires good starting values in order to reach a globally optimal solution. In Chapters 4 and 6 the author extends an idea of Posse (2001) that draws on graph theoretic approaches to clustering. In particular, Posse suggests that the model-based clustering can be initialised from a clustering generated from the minimum spanning tree (MST) of the data. The particular extension of this procedure proposed by the author is to use the MST to determine observations that should initially be allocated to the noise category.

Despite the versatility of model-based clustering, the technique imposes some considerable parametric assumptions on the shape of clusters. Recent developments in clustering using graph theoretic approaches based on the MST make no such assumptions (Xu et al., 2002; Xu et al., 2003). One future line of research that the author would like to pursue, is a comparison of model-based clustering with these non-parametric approaches. In particular, the clustering approaches could be compared in their ability to define partitions of the data that isolated the greatest differences in the HPF.

3.ii. Estimation of the HPF

Of all the different components of a hedonic pricing study, it is the econometric estimation of the HPF that has received the most attention in the literature. The research reported in this thesis has endeavoured to add to that substantial literature. In particular, the author has been concerned with introducing flexibility into the specification of the HPF and with issues relating to spatial autocorrelation (SA) in regression residuals.

As reported above, methods of data partitioning have been employed as a means of isolating areas of the HPF with substantially different slopes. Furthermore, and as described in Chapter 5, a semiparametric estimator (Robinson, 1988) is used to introduce further flexibility into the specification of the HPF. Improvements are made on previous applications of this model by allowing both selected property characteristics and the influence of location to enter the econometric model nonparametrically.

As might be anticipated, analysis of the regression residuals from the hedonic price regressions reveals evidence of SA. In Chapter 5 the author employs an estimation strategy proposed by Kelejian and Prucha (1999) which specifically accounts for SA and returns robust estimates of the model's parameters and their variance. As far as the author is aware, this is the first application to combine semiparametric methods with the Kelejian and Prucha GMM estimator to provide robust estimates of the parameters of a HPF.

Chapter 7 re-examines the issue of SA in regression residuals and provides possibly the most innovative contributions of this thesis. Specifically, Chapter 7 describes an improved procedure for examining the pattern of SA in regression residuals and proposes a new approach to the estimation of econometric models showing evidence of SA.

Common to all tests of SA is the need for researchers to specify *a priori* the spatial area over which to examine residuals for evidence of correlation. Typically, this is done by assuming that only the residuals of observations within a certain distance of each other will be correlated. Presently, however, there is no established procedure for determining this distance. The first novel contribution provided in Chapter 7 is to propose a procedure in which the *spatial correlogram* constructed using Moran's *I* statistic is used to establish the pattern of spatial dependence in regression residuals.

Of course, evidence of SA in regression residuals is evidence of spatial features not captured in the specification of the model. Commonly researchers assume that such features comprise the many subtle nuances of location that could adequately be handled by modelling the nuisance process. Models that follow this line of reasoning are called *Spatial Error Dependence* (SED) models. Alternatively, the researcher might believe that the unobserved features of the spatial environment are substantial and that their absence from the model is likely to induce missing variable bias in the parameter estimates. In this case one way forward is to account for the missing covariates by spatially smoothing the data using a nonparametric kernel procedure that Gibbons and Machin (2003) term a *Smooth Spatial Effects* (SSE) estimator.

Application of the SSE requires the researcher to determine a spatial smoothing bandwidth, that is, to stipulate the spatial area over the data should be smoothed. Previously the choice of smoothing bandwidth for the SSE estimator has been either ad hoc (Gibbons and Machin, 2003; Gibbons, 2003) or has used a data-driven *cross-validation* procedure (Clapp, 2004) that can be shown to be non-optimal in this case (Opsomer, 2001).

The second major contribution of this paper is to advocate an alternative procedure for selecting the smoothing bandwidth that specifically recognises the spatial nature of the problem. The proposed procedure involves repeated estimation of the SSE model using progressively larger bandwidths. At each iteration, Moran's I statistic is calculated to assess the degree of SA in the residuals over the spatial scale identified by the correlogram. The optimal bandwidth is selected as that bandwidth at which the computed value of Moran's I matches its expectation under the hypothesis of spatially uncorrelated residuals. At this bandwidth, the test statistic indicates that spatial smoothing of the data has accounted for the influence of omitted locational covariates. This bandwidth selection procedure is an application of Ellner and Seifu's (2002) *Residual Spatial Autocorrelation* (RSA) criterion. Unlike other bandwidth selection procedures, the RSA criterion is both an intuitive procedure and, through the elimination of SA, permits statistical inference and testing of the model to proceed using standard econometric tools.

Using the RSA criterion to select smoothing bandwidth, the SSE model and SED model are applied to the Birmingham data. For certain partitions of the data, the parameter estimates are found to differ significantly. In these cases it is clear that applying the standard SED models is incorrect since the parameter estimates from this model will suffer from omitted locational covariate bias.

The RSA procedure is extremely computer-intensive requiring repeated estimation of the SSE at various bandwidths. As such an additional achievement of the research carried out in this thesis has been the development of procedures written in a combination of Gauss and Fortran to allow for rapid estimation of the SSE. These procedures make use of sparse matrices to summarize the spatial relationships between observations and allow fast implementation of local linear regression by linearly binning the data and employing FFT-based calculations. The procedures were found to be many hundreds of times faster than naïve implementations of local linear regression.

One drawback of applying the RSA criterion to the selection of spatial smoothing bandwidth, is that it imposes isotrophy. In other words, the procedure assumes that the same smoothing bandwidth can be used to account for missing locational covariates over the entire urban area. In reality, of course, the spatial influence of missing covariates will vary considerably over space; some locations will be characterised by missing covariates that influence property prices over a large area, in other locations their influence may be highly localised.

In future research the author intends extending and refining the RSA procedure. In particular, it is envisaged that the pattern of SA pertaining to each individual observation could be captured by constructing a spatial correlogram using local indicators of SA (LISAs) such as local Moran's *I*. The significance of SA at increasing distance from each observation location could be determined using permutation tests. Accordingly an initial bandwidth for each observation could be chosen as the distance from each observation at which it's LISA statistic became insignificant. Assuming that the optimal bandwidth for each observation is proportional to this distance, the problem would be to find the value of a parameter that when multiplied with each observation's intial bandwidth gives the optimal set of smoothing bandwidths for the SSE. Again, it should be possible to employ Moran's *I* statistic for this purpose. Specifically, it is well-known that summing Local Moran's *I* for each observation gives the value of Global Moran's *I*. The optimal value of the multiplicative parameter could be taken as the value at which Global Moran's *I* reached its expected value under a random distribution of errors across space.

3.iii. Estimation of the demand functions

The final part of this thesis concerns itself with the estimation of (mythical) demand functions for 'peace and quiet' from traffic-related noise. In pursuit of this goal, the author has endeavoured to maintain the same standards of methodological rigour achieved elsewhere in this thesis. Indeed, the analysis has required the application of numerous sophisticated econometric tools. However, the undertaking has been complicated and occasionally compromised by the limitations of the data set available to the author. This section reviews and critiques some of the key assumptions and the methodological techniques used in the estimation of the demand functions for 'peace and quiet'.

Estimation of the demand relationship depends crucially on having data that identifies how much 'peace and quiet' a household with particular characteristics will choose when faced by different prices. Accordingly, in Part 3 of this thesis, it is assumed that by partitioning the Birmingham data we are identifying different market segments each characterised by an independent pricing structure.[58] This identification strategy depends on assuming that there is sufficient overlap in the socioeconomic characteristics of these market segments to observe households with similar characteristics participating in markets with different prices for noise.

Unfortunately, there is some circularity in this estimation strategy. In order to identify the demand relationship, the data must be divided into different market segments. The market segments are defined in part by the socioeconomic characteristics of households. Consequently the market segments define relatively distinct socioeconomic groupings. Accordingly, the degree of overlap between market segments may perhaps be rather less than might be hoped. Rather better would be to repeat the first stage analysis for a second urban area characterised by a similar socioeconomic mix of inhabitants as that in Birmingham. The combination of the output from that second study with the Birmingham study would bring together data from similar households choosing in distinct markets and would provide a far stronger basis for identification of the demand relationship. Nevertheless, the analysis continues under the assumption that the Birmingham data provides sufficient variation in the price of 'peace and quiet' and sufficient overlap in the characteristics of households participating in different market segments to allow identification of the demand relationship.

The author has also faced a number of challenges in overcoming specific limitations of the data. The first major issue concerns the fact that the dataset lacks information on the particular characteristics of the households making the property purchases. Since both household income and household socioeconomic characteristics are

[58] As discussed in detail in Section 2 of this Chapter, this is not the only interpretation that might be given to the observation that the different partitions of the data show significantly different price structures.

arguments in the demand relationship, this is a considerable shortcoming of the Birmingham dataset.

To overcome the absence of income information, a two-stage budgeting hypothesis is invoked. Under this rational simplifying assumption, the relevant argument in the demand relationship is not household income, but the household's total expenditure in purchasing a property. Of course, this information is available since the property purchase prices are known.

There is less room for manoeuvre with regards to the specific socioeconomic characteristics of households. Lacking any alternative, the average socioeconomic characteristics of the households' neighbourhoods are included as a crude proxy for households' actual characteristics.

The second data issue concerns the fact that the prices estimated in the first stage of the hedonic analysis are one-off payments equivalent to the prices paid for durable goods. Rather, for the purposes of demand analysis, it would be preferable to have prices expressed per period, say annually. Of course, in the property market such annual payments are known as property rents.

Under an assumption of perfectly operating property markets, the key issue in converting purchase prices into annual payments is determining the rate of discount used by households when comparing property prices with property rents. In Chapter 9, this rate of discount is estimated using a small dataset relating property sales prices to property rental prices in Birmingham in 2003. Some reasonably defendable assumptions concerning purchaser myopia and the liquidity of property markets are required in order to transfer this implied rate of discount back to 1997. An area worthy of further investigation would be to determine the sensitivity of the research findings to changes in the assumed discount rate.

A third major data issue concerns the fact that the first stage hedonic regressions return results that allow for negative prices for 'peace and quiet'. Clearly, negative prices defy both theory and common sense. For aircraft noise, there are several reasons why such parameter estimates might be considered unreliable. Accordingly no demand function is estimated for aircraft noise. For road traffic noise, on the other hand, the parameters are considered reasonably robust. In this case, the negative prices that are observed in two of the seven market segments are assumed to result from sampling error. That is, it is assumed that the 'true' implicit prices for 'peace and quiet' from road traffic noise are in the region of zero such that negative prices can be set to a value of zero in the regression analysis.

Our econometric analysis seeks to estimate two separate demand equations; one for 'peace and quiet' from road traffic noise and a second for 'peace and quiet' from rail traffic noise. As in all econometric analyses, the author is forced to impose assumptions on the econometric specification of these demand relationships.

In theory, the implicit prices of all property attributes should enter the demand relationship for 'peace and quiet'. In practice this would require the inclusion of an impractically large number of covariates. Consequently, cross-price effects for all attributes other than those relating to the noise environment are ignored. Moreover, to maintain simplicity the two demand relationships are estimated as a linear demand system. In matter or fact, economic theory implies numerous relationships between the variables and parameters of the demand system that are contravened by the linear specification. Several econometric specifications have been developed that seek to explicitly impose the restrictions implied by economic theory, and the author once again acknowledges that applying these specifications to the Birmingham data may be worthy of further investigation.

One advantage of the linear regression model is that within that framework, procedures to account for the endogeneity of prices and expenditure and for the censoring of the dependent variable are well-developed. To account for censoring of the dependent variable (i.e. households choosing properties unaffected by road and/or rail traffic noise) the author adopts a Tobit estimator. Whilst the form of this estimator is convenient, it imposes fairly strict distributional assumptions on the model. To account for endogeneity (i.e. simultaneity in the choice of quantities, prices and expenditure) the author employs an instrumental variables procedure. The instrumental variables themselves have been chosen under the assumption that uncaptured elements of the joint choice process reflect specific characteristics of the household and their attitude to noise. As such, the author follows Cheshire and Sheppard (1998) and selects as instrumental variables the prices and expenditure chosen by other nearby households. Here "nearby" is defined as proximity in a multidimensional space defined by geographic location, socioeconomic characteristics and property characteristics. Whilst the author has carried out tests to confirm that prices and expenditure are not exogenous in the demand functions, no similar test has been carried out to confirm the validity of the instrumental variables. Further investigation of the theoretical suitability and empirical validity of the chosen instrumental variables would be appropriate.

The econometric model used to estimate the (mythical) demand functions for rail and road noise can be described as an IV Tobit estimator. In arriving at this estimator and the data to which it is applied, the author has been forced to make numerous assumptions. In Chapter 9, each of these assumptions has been highlighted and discussed in detail. In each case arguments have been forwarded to justify modelling decisions. The author believes that, despite the limitations imposed by the data, the resultant analysis maintains a high degree of methodological rigour and integrity.

4. Policy-Relevant Outputs

The research reported in this thesis has generated two sets of policy-relevant outputs. First the parameters of the HPFs estimated in the first stage regressions, provide an indication of the impact on property prices brought about by changes in their exposure to traffic-related noise. Such values may be of relevance to compensation claims made in response to public works on roads or airports. According to the Land Compensation Act (1973), compensation for public works should be made according to "the difference between: (i) the price a purchaser would pay for the property with the public works in use but with the physical factors no worse than they were before the schemes, and (ii) the price a purchaser would pay with the public works in use with the present or anticipated effect of the physical factors" (para 14.25). In this case, the "physical factors" refer to elements of transport-related disturbance including noise, vibration and pollution. Accordingly, the parameters of the HPF, which record the change in property prices brought about by changes in exposure to traffic-related noise, could be employed convincingly in directing the level of such compensation awards.

In this study, it has been assumed that traffic-related noise exposure only impacts on property prices once it exceeds 55dB L_{EQ}. The 55dB L_{EQ} threshold reflects the level of urban background noise as defined by the UK Department for Transport in their transport appraisal guidance (www.webTAG.org.uk para 1.3.7 of unit 3.3.2). With regards to the coefficients on the noise pollution variables, the HPFs estimated for the Birmingham data are generally pleasing. In the main, the coefficient estimates are of the correct sign and mostly have plausible magnitudes. The models reveal that a 1dB increase in road noise reduces property prices by between 0.21% and 0.53%, depending on market segment. The rail noise estimates are relatively larger in magnitude, indicating that on average a 1dB increase in rail noise will reduce property prices by 0.67%. Little evidence is found of a relationship between air traffic noise and property prices. This finding is almost certainly a result of the econometric approach which tends to subsume wide-area spatial effects into other parameters of the model.

As explicitly discussed in Part 1 of this thesis, however, the responsiveness of property prices to changes in exposure to traffic-related noise will be determined by the unique conditions of property supply and demand that exist in any particular market. As such, coefficients from the HPF of the Birmingham property market are only strictly applicable to compensation claims being made in that market.

The second policy-relevant output from this thesis comes from the second stage regressions in which mythical demand functions for 'peace and quiet' are estimated. The key advantage of identifying demand functions is that these provide a means of calculating valid measures of welfare change from changes in 'peace and quiet'. Indeed, it is these measures of welfare change, rather than the price changes identified in the first stage regressions, that would be of relevance to policy-makers using cost-benefit analysis to evaluate projects changing the noise environment

It should be noted that there are some limitations to the comprehensiveness of these welfare measures. In particular, the values only record the welfare changes experienced by the residents of an area witnessing an environmental change, not those experienced by those who work or travel through such an area. Furthermore, and as noted in Section 2 of this Chapter, these values do not account for the wider market readjustments that might be precipitated by environmental change.

There are other techniques that might be used to calculate welfare values for changes in environmental quality; most notably stated preference techniques such as contingent valuation. However, these approaches suffer from the same limitations as those just identified for the hedonic pricing method plus numerous others unique to these techniques. Moreover, the great advantage of the hedonic pricing method is that it is based on evidence from actual market behaviour. That is, the values from the hedonic method are based on actual choices involving actual purchases in real markets. Of course, this is also one drawback of the hedonic pricing method as implemented in this research. In particular the data refer to property purchases and as such the results can only apply to those households that participate in the property market.

The welfare estimates calculated from the mythical demand curves are values per annum reported in 1997 prices. For road noise these range from a low of £19 per annum for a low-income household experiencing a 1dB change in a relatively quiet environment (56dB), to a high of £105 per annum for a high-income household experiencing a 1dB change in a noisy environment (80dB). For rail noise the values are somewhat higher ranging from £52 per annum for a low-income household experiencing a 1dB change in a relatively quiet environment (56dB) to £178 per annum for a high-income household experiencing a 1dB change in a noisy environment (80dB).

Despite the considerable complexity of the task, this thesis has been successful in its endeavour to estimate demand relationships for 'peace and quiet'. Unfortunately, such is the pioneering nature of this research, there are no equivalent studies with which the output from this thesis can be compared. Nonetheless, the welfare estimates derived from these demand functions pass a crucial, if not highly technical test of validity; they appear to be of a reasonable magnitude.

Finally, the author has discussed how the welfare estimates from the Birmingham study might be transferred to households living in other locations. In transferring the demand function from one location to another it must be assumed that the basic structure of preferences remains the same between households residing in Birmingham and the transfer location. Unfortunately, such an assumption is only testable if a similar analysis to that carried out in the Birmingham study were carried out for other locations.

In Chapter 9, the author describes a pragmatic method by which welfare values from this study might be transferred to other locations. This transfer procedure attempts to adjust for differences in the wealth of inhabitants living in different locations and to account for problems with property price inflation that might complicate the transfer of values across time as well as across locations. Whilst these procedures are open to criticism, the author argues that they represent an operationally pragmatic means of carrying out benefits transfer exercises.

5. Final Comments

The research reported in this thesis records one of the most comprehensive hedonic pricing studies carried out to date. The research is based upon one of the richest property market datasets ever constructed. This dataset comprises over 10,000 observations of property purchases in the City of Birmingham in 1997. Moreover, each observation of a property sale is accompanied by an array of accurate details of the characteristics of that property.

The majority of previous studies have attempted simply to identify the implicit price function for 'peace and quiet' using property market data. Since that implicit price function is simply a reflection of the specific conditions of supply and demand that exist in a property market, these studies provide no basis for calculating transferable welfare values.

In this research, the author has attempted to go one step further and estimate demand functions for 'peace and quiet'. Since these functions identify a household's underlying preferences, that is, how they are prepared to trade-off between money and 'peace and quiet', they can be used as objects for transferring benefits across locations.

The discussion in this Chapter has highlighted the major theoretical and methodological challenges addressed in this thesis and commented on the degree to which these have been successfully overcome. Furthermore this Chapter has indicated those lines of research emanating from this work worthy of further investigation.

With regards to estimation of the HPF, this thesis reports on new techniques relating to data partitioning and spatial smoothing that show considerable potential as methods of data analysis. In future research, the author intends investigating these methods further.

With regards to the estimation of demand functions, it is the author's opinion that progress can only be achieved through the collation of data sets that combine property market data with data collected directly from purchasing households, preferably in several distinct urban areas. The author would relish the renewed challenge of undertaking such a study.

APPENDICES

APPENDIX A: ANALYTICAL MODELS OF EQUILIBRIUM IN HEDONIC MARKETS

A1. The Normal-Linear-Quadratic Model: A simple closed-form solution

A number of researchers (e.g. Tinbergen, 1956; Epple, 1987; Tauchen and Witte, 2001; Ekeland et al., 2002) have investigated a particularly simple form of the Rosen model for hedonic market equilibrium that results in a closed form solution to Equation (29) of Chapter 1. This model has been labelled the *Normal-Linear-Quadratic* (NLQ) model by Ekeland et al. (2002).

Whilst imposing reasonably restrictive assumptions concerning the various behavioural functions that determine household and landlord behaviour, the NLQ model provides valuable insights into the workings of hedonic markets. Further, an understanding of this simple model makes some of the econometric issues to be discussed in Parts 2 and 3 of this thesis, far more transparent.

Here we roughly follow the exposition of Tauchen and Witte (2001) and Ekeland et. al (2002). Let us assume first that the utility function (Equation 5a) of Chapter 1 takes the simple quadratic form;

$$U(z,x;A,\boldsymbol{\alpha}) = \boldsymbol{\alpha}'z - \frac{1}{2}z'Az + x \qquad \text{(A1)}$$

Here $\boldsymbol{\alpha}$, the vector describing the particular tastes of the household (and assumed to be a function of household characteristics, s), makes up the linear terms of the utility function and is assumed to have the same dimensions as z (that is, $\boldsymbol{\alpha}$ contains K elements). Whilst A, the matrix of commonly held parameters of the utility function, is assumed to be symmetric and positive definite with dimensions K by K.

The quadratic functional form is commensurate with diminishing marginal utility in z and gives rise to indifference curves that take on the classic, concave to the origin shape depicted in Figure 3. Moreover, the assumption that x (money to spend on other goods) enters the utility function as a simple additive term implies that the indifference curves (and hence bid curves) will be vertically parallel. This form of utility function is called *quasilinear* and proves to have some interesting properties that we shall discuss in more detail later.

Now, the utility provided by a property with characteristics z to a household with income y and characteristics $\boldsymbol{\alpha}$ is given by;

$$\boldsymbol{\alpha}'z - \frac{1}{2}z'\boldsymbol{A}z + y - P(z;\boldsymbol{\Gamma},\gamma) \quad \text{(A2)}$$

The household seeks a property whose characteristics, z, maximise this function such that the first order conditions for an optimal residential choice are given by;

$$\boldsymbol{p}_z(z;\boldsymbol{\Gamma},\gamma) = \boldsymbol{\alpha} - \boldsymbol{A}z \quad \text{(A3)}$$

where $\boldsymbol{p}_z$ is the vector of implicit prices with typical element p_{z_i}. For now we assume that every household's second order condition is globally satisfied. After solving for the equilibrium HPF we shall check to see if this is true.

The set of K equations in (A3) describe the implicit prices a household with preferences given by $\boldsymbol{\alpha}$ and $\boldsymbol{A}$ would be willing to pay in order to obtain a given vector of property characteristics. Hence, given a schedule of implicit prices, Equation (A3) completely solves the household's choice problem. Indeed, Equation (A3) defines the set of *inverse ordinary demand functions*, which we denote;

$$\boldsymbol{b}^d(z;y,\boldsymbol{\alpha},\boldsymbol{A}) = \boldsymbol{\alpha} - \boldsymbol{A}z \quad \text{(A4)}$$

Our assumed functional form for utility ensures that the inverse ordinary demand functions take on their expected downward-sloping appearance. Moreover, the demand curves are linear which accounts for the "linear" in the name of the NLQ model.

Alternatively, we could solve Equation (A3) for z in order to derive the *ordinary demand functions*, $z^d(\boldsymbol{A},\boldsymbol{\alpha},\boldsymbol{\Gamma},\gamma,y)$. Of course to do so requires knowledge of the functional form of the HPF, since ordinary demand functions, unlike the inverse functions, are dependent on market prices. Notice that with quasilinear preferences, the inverse ordinary demand functions (and as we shall see shortly, the ordinary demand functions) are actually independent of income; households with taste parameters $\boldsymbol{\alpha}$ will choose the same vector of marginal prices/property characteristics regardless of their income.

An alternative approach to solving the household's problem is to follow the exposition in Section 4 of this Chapter and derive the *bid function*. Inverting the utility function (Equation A1) provides the *indifference function*;

$$x(z;u,\boldsymbol{\alpha},A) = u - \boldsymbol{\alpha}'z + \frac{1}{2}z'Az \qquad \text{(A5)}$$

from which we formulate the *bid function* according to;

$$\theta(z;\, y,u,\boldsymbol{\alpha},A) = y - \left(u - \boldsymbol{\alpha}'z + \frac{1}{2}z'Az\right) \qquad \text{(A6)}$$

Taking derivatives of Equation (A6) with respect to the choice vector z reveals the *marginal bid functions*;

$$\boldsymbol{b}^{d}(z;u,\boldsymbol{\alpha},A) = \boldsymbol{\alpha} - Az \qquad \text{(A7)}$$

which are identical to the inverse ordinary demand functions of Equation (A3). Indeed, it turns out that the marginal bid function is simply an alternative approach to constructing the *inverse compensated demand function.*[59] In the special case of quasilinear preferences, utility drops out of the marginal bid function and the two forms of inverse demand function are identical.

Of course, if we had chosen an alternative formulation for household preferences in which x (money to spend on other goods) was not a simple additive term in the utility function, then this identity would not hold. In that case, income would be an argument in the ordinary demand functions and utility an argument in the

[59] Alternatively this function can be derived by solving the dual of the household's residential choice problem; that is to minimise total expenditure so as to achieve a specified level of utility.

compensated demand functions and the two functions would describe different relationships between prices and quantities.

On the landlords side of the market let us assume that each landlord lets out one property with the quadratic cost function;

$$C(z,\underline{z},\boldsymbol{\beta},\boldsymbol{B}) = \boldsymbol{\beta}'(z-\underline{z}) + \frac{1}{2}(z-\underline{z})'\boldsymbol{B}(z-\underline{z}) \tag{A8}$$

Here, $\boldsymbol{B}$ and $\boldsymbol{\beta}$ are the parameters of the cost function describing input prices and the technologies available to landlords. $\boldsymbol{B}$ is assumed to be a symmetric, positive definite, K by K matrix that is identical for all landlords. Whilst $\boldsymbol{\beta}$ is a K vector of cost parameters that vary across landlords. Notice that the function described by Equation (A8) implies that landlords' costs in providing a property with characteristics z, are determined only by that part of the property attributes that are not supplied costlessly to the landlord.[60] Furthermore, the quadratic functional form implies that marginal costs increase as landlords make adaptations to their properties characteristics that take them further and further from the costlessly provided $\underline{z}$. Increasing marginal costs is commensurate with the upward sloping offer curves depicted in Figure 6.

Since the profits from providing a property with characteristics z are given by the identity $\pi \equiv P(z) - C(z)$, the first order conditions defining the optimal choice for a landlord are;

$$\boldsymbol{p}_z(z;\boldsymbol{\Gamma},\gamma) = \boldsymbol{\beta} + \boldsymbol{B}(z-\underline{z}) \tag{A9}$$

Again we assume that the second order conditions are globally satisfied and will check the veracity of this assumption once we have solved for the equilibrium HPF.

The set of K equations in (A9) specify the marginal prices a landlord would require in order to supply a particular vector of property characteristics. Thus Equation (A9) defines the *inverse supply functions*;

---

[60] The model presented here then is a simple generalisation of that in Tauchen and Witte (2001) and Ekeland et. al (2002).

$$\boldsymbol{b}^s(z;\underline{z},\boldsymbol{\beta},\boldsymbol{B}) = \boldsymbol{\beta} + \boldsymbol{B}(z-\underline{z}) \tag{A10}$$

Again our assumptions concerning the cost function ensure that the supply functions take on the expected, upward-sloping appearance.

We could solve Equation (A9) to form the more familiar *supply functions*. However, the dependence of implicit prices on z means that these take the as yet undetermined form, $z^s = z^s(\underline{z},\boldsymbol{B},\boldsymbol{\beta},\boldsymbol{\Gamma},\gamma)$.

Once again, these K equations could be derived by following the exposition in Section 5 of this Chapter and first forming the *offer function* (Equation 18), which in this case would be written;

$$\phi(z;\underline{z},\boldsymbol{\beta},\boldsymbol{B},\pi) = \pi + C(z,\underline{z},\boldsymbol{\beta},\boldsymbol{B}) \tag{A11}$$

Notice that the partial derivatives of the *offer function* with respect to each property characteristic gives a set of K equations identical to the right-hand side of Equation (A10); the *inverse supply functions* and *marginal offer functions* are one in the same. Unlike the household's problem the two will be identical irrespective of the functional form of the cost function.

Establishing the equilibrium HPF in this market requires making assumptions concerning how preference and cost parameters ($\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ respectively) are distributed across households and landlords in the property market. The "Normal" part of the NLQ model refers to the assumptions that;

- the vector $\boldsymbol{\alpha}$ follows a multivariate normal distribution with mean $\bar{\boldsymbol{\alpha}}$ and variance $\boldsymbol{V_\alpha}$
- the vector $\boldsymbol{\beta}$ follows a multivariate normal distribution with mean $\bar{\boldsymbol{\beta}}$ and variance $\boldsymbol{V_\beta}$

Furthermore, let us assume that the costlessly supplied attributes $\underline{z}$ are randomly distributed across properties in the market. In this model, therefore, we assume that each landlord is randomly allocated a property with an initial set of attributes ($\hat{z}$ in the terminology of the last section). Exogenous changes in the locational, neighbourhood and/or environmental characteristics of a property might result in changes to the levels of costlessly supplied attributes. Let us assume that these changes are independent of the property characteristic vector z, or the parameters of the landlords cost function, $\boldsymbol{B}$ and $\boldsymbol{\beta}$, and that;

- the vector $\underline{z}$ follows a multivariate normal distribution with mean $\bar{\underline{z}}$ and variance $\boldsymbol{V}_{\underline{z}}$

Given these distributions we could use the change of variable technique to establish the demand and supply densities. Equating these densities (as per Equation 29) provides an equation that can be solved for the market clearing HPF $P(z;\boldsymbol{\Gamma},\gamma)$. In the NLQ model, it transpires that a quadratic price function of the form;

$$P(z;\boldsymbol{\Gamma},\gamma)=\gamma'z+\frac{1}{2}z'\boldsymbol{\Gamma}z \tag{A12}$$

satisfies the conditions for an equilibrium, hence the term "Quadratic" in the NLQ model .

Again following Tauchen and Witte (2001) and Ekeland et. al (2002), we can begin by assuming a quadratic HPF of the form given by Equation (A12), then check to see that this is indeed a correct solution. At the same time it will prove beneficial for later discussion to determine how the price parameters (the K-vector, γ and the K by K matrix, $\boldsymbol{\Gamma}$) are related to the underlying structural parameters of the utility and cost functions.

Armed with a functional form for the hedonic price equation (A12), we derive the implicit prices;

$$\boldsymbol{p}_z(z;\boldsymbol{\Gamma},\gamma)=\gamma+\boldsymbol{\Gamma}z \tag{A13}$$

and the demand and supply functions can now be seen to take the explicit forms;

$$z^d=(\boldsymbol{\Gamma}+\boldsymbol{A})^{-1}(\alpha-\gamma) \tag{A14}$$

$$z^s = (\boldsymbol{\Gamma} - \boldsymbol{B})^{-1}(\boldsymbol{\beta} - \boldsymbol{B}\underline{z} - \boldsymbol{\gamma}) \tag{A15}$$

Now since $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\underline{z}$, the normally distributed parameters of the model, enter these equations linearly, we can conclude that the quantities of property characteristics demanded and supplied in the market are also normally distributed. Consequently, equating demand with supply reduces to the relatively simple task of equating the mean of demand with the mean of supply and the variance of demand with the variance of supply. The first of these relationships, equality of means, is given by the expression;

$$(\boldsymbol{\Gamma} + \boldsymbol{A})^{-1}(\bar{\alpha} - \boldsymbol{\gamma}) = (\boldsymbol{\Gamma} - \boldsymbol{B})^{-1}(\bar{\beta} - \boldsymbol{B}\bar{\underline{z}} - \boldsymbol{\gamma}) \tag{A16}$$

whilst the second of these relationships, equality of variances, provides the expression;

$$(\boldsymbol{\Gamma} + \boldsymbol{A})^{-1}\boldsymbol{V}_{\alpha}(\boldsymbol{\Gamma} + \boldsymbol{A})^{-1} = (\boldsymbol{\Gamma} - \boldsymbol{B})^{-1}\left(\boldsymbol{V}_{\beta} + \boldsymbol{B}\boldsymbol{V}_{\underline{z}}\boldsymbol{B}\right)(\boldsymbol{\Gamma} - \boldsymbol{B})^{-1} \tag{A17}$$

where the symmetry of $\boldsymbol{V}_{\alpha}, \boldsymbol{V}_{\beta}, \boldsymbol{V}_{\underline{z}}, \boldsymbol{A}$ and $\boldsymbol{B}$ allows us to avoid the transpose notation and our assumption of independence between $\boldsymbol{\beta}$ and $\underline{z}$ removes the need for an element reflecting the covariance of these two variables.

In general, therefore, Equation (A17) can be solved to give an expression for $\boldsymbol{\Gamma}$ in terms of the known parameters. To construct such a solution requires forming the square roots of the matrices $\boldsymbol{V}_{\alpha}$ and $\boldsymbol{V}_{\beta} + \boldsymbol{B}\boldsymbol{V}_{\underline{z}}\boldsymbol{B}$. Since, $\boldsymbol{V}_{\alpha}$, $\boldsymbol{V}_{\beta}$ and $\boldsymbol{V}_{\underline{z}}$ are variance matrices they are both symmetric and positive definite. Likewise we have assumed that $\boldsymbol{B}$ is a symmetric and positive definite matrix. As a result, both $\boldsymbol{V}_{\alpha}$ and $\boldsymbol{V}_{\beta} + \boldsymbol{B}\boldsymbol{V}_{\underline{z}}\boldsymbol{B}$ will themselves by symmetric and positive definite[61] and hence have non-singular square roots. We denote these square root matrices, $\boldsymbol{S}_{\alpha}$ and $\boldsymbol{S}_{\beta\underline{z}}$ such that by definition $\boldsymbol{S}_{\alpha}\boldsymbol{S}'_{\alpha} = \boldsymbol{V}_{\alpha}$ and $\boldsymbol{S}_{\beta\underline{z}}\boldsymbol{S}'_{\beta\underline{z}} = \boldsymbol{V}_{\beta} + \boldsymbol{B}\boldsymbol{V}_{\underline{z}}\boldsymbol{B}$. Of course, there are likely

[61] For proofs of the properties of positive definite matrices see Johnson (1970) and Greene (1993).

to be numerous matrices that conform to these definitions, though, as we shall see shortly, we can narrow down possible candidates by enforcing the second order conditions.

Armed with the square root matrices $\boldsymbol{S}_{\alpha}$ and $\boldsymbol{S}_{\beta\underline{z}}$, Equation (A17) can be solved to give;

$$\boldsymbol{\Gamma}_c = \left(\boldsymbol{I} - \boldsymbol{S}_{\alpha}\boldsymbol{S}_{\beta\underline{z}}^{-1}\right)^{-1}\left(\boldsymbol{B} + \boldsymbol{S}_{\alpha}\boldsymbol{S}_{\beta\underline{z}}^{-1}\boldsymbol{A}\right) \tag{A18}$$

where $\boldsymbol{\Gamma}_c$ represents one of a number of candidate solutions for $\boldsymbol{\Gamma}$ depending on the choice of $\boldsymbol{S}_{\alpha}$ and $\boldsymbol{S}_{\beta\underline{z}}$. For $\boldsymbol{\Gamma}_c$ to be an actual solution, it must be symmetric. We can eliminate further candidate solutions by applying the second order conditions for optimal choices by households and landlords. Inserting the specific implicit prices from Equation (A13) in the first order conditions (Equation A3 for households and Equation A9 for landlords) reveals that the second order conditions require $-(\boldsymbol{\Gamma} + \boldsymbol{A})$ and $(\boldsymbol{\Gamma} - \boldsymbol{B})$ to be negative definite. As a consequence a solution requires that;

$$\begin{aligned} &\left(\boldsymbol{I} - \boldsymbol{S}_{\alpha}\boldsymbol{S}_{\beta\underline{z}}^{-1}\right)^{-1}\left(\boldsymbol{B} + \boldsymbol{S}_{\alpha}\boldsymbol{S}_{\beta\underline{z}}^{-1}\boldsymbol{A}\right) + \boldsymbol{A} \\ &= \left(\boldsymbol{I} - \boldsymbol{S}_{\alpha}\boldsymbol{S}_{\beta\underline{z}}^{-1}\right)^{-1}\left(\boldsymbol{A} + \boldsymbol{B}\right) \end{aligned} \tag{A19}$$

and

$$\begin{aligned} &\left(\boldsymbol{I} - \boldsymbol{S}_{\alpha}\boldsymbol{S}_{\beta\underline{z}}^{-1}\right)^{-1}\left(-\boldsymbol{B} - \boldsymbol{S}_{\alpha}\boldsymbol{S}_{\beta\underline{z}}^{-1}\boldsymbol{A}\right) + \boldsymbol{B} \\ &= -\left(\boldsymbol{I} - \boldsymbol{S}_{\alpha}\boldsymbol{S}_{\beta\underline{z}}^{-1}\right)^{-1}\boldsymbol{S}_{\alpha}\boldsymbol{S}_{\beta\underline{z}}^{-1}\left(\boldsymbol{A} + \boldsymbol{B}\right) \end{aligned} \tag{A20}$$

be positive definite. Although solutions conforming to Equations (A19) and (A20) do not necessarily exist for all parameter values, there are open subsets of parameters for which they do (Tauchen and Witte; 2001).

Having solved for Γ, the equality of means condition (Equation A16) provides a solution for γ;

$$\gamma = \left[(\boldsymbol{\Gamma} - \boldsymbol{B})^{-1} - (\boldsymbol{\Gamma} + \boldsymbol{A})^{-1}\right]^{-1} \left[(\boldsymbol{\Gamma} - \boldsymbol{B})^{-1} \left(\bar{\boldsymbol{\beta}} - \boldsymbol{B}\bar{\underline{z}}\right) - (\boldsymbol{\Gamma} + \boldsymbol{A})^{-1} \bar{\boldsymbol{\alpha}}\right] \tag{A21}$$

The solutions given by Equation (A18) and Equation (A21) provide an insight into the nature of equilibrium in the property market. The parameters of the HPF are determined by the parameters of the utility and cost functions and how these are distributed across households and landlords in the urban area. Also, changes in the distribution of exogenously provided property attributes, $\underline{z}$, will directly influence the equilibrium HPF. For example, the model predicts that government interventions to improve environmental quality will have direct but complex implications for the equilibrium HPF through changes in the value of $\bar{\underline{z}}$ and $\boldsymbol{S}_{\boldsymbol{\beta}\underline{z}}$.

To further understand the insights proffered by the NLQ model let us examine some special cases of the general solution. First, assume that the random variables describing the parameters of households' preferences, $\boldsymbol{\alpha}$, are independent of each other. Likewise, make the same assumption for the parameters of landlords' costs, $\boldsymbol{\beta}$, and the exogenously provided attributes for each property, $\underline{z}$. In this case, the variance matrices $\boldsymbol{V}_{\boldsymbol{\alpha}}$, $\boldsymbol{V}_{\boldsymbol{\beta}}$ and $\boldsymbol{V}_{\underline{z}}$ will be diagonal. In a similar vein, assume that $\boldsymbol{A}$ and $\boldsymbol{B}$ are diagonal matrices, such that the utility and cost functions do not allow for interactions effects between the levels of property attributes. Epple (1987), Tauchen and Witte (2001) and Ekeland et al. (2003) all discuss this simplest case of the model.

As shown by Epple (1987) and reiterated by Tauchen and Witte (2001), the independent case leads to a particularly simple solution. First construct the diagonal matrices $\boldsymbol{V}_{\boldsymbol{\alpha}}^{1/2}$ and $\boldsymbol{V}_{\boldsymbol{\beta}\underline{z}}^{1/2}$ whose diagonal elements consist of the positive square root of the corresponding element in the matrices $\boldsymbol{V}_{\boldsymbol{\alpha}}$ and $\boldsymbol{V}_{\boldsymbol{\beta}} + \boldsymbol{B}\boldsymbol{V}_{\underline{z}}\boldsymbol{B}$. Of course, the negative of the matrices $\boldsymbol{V}_{\boldsymbol{\alpha}}^{1/2}$ and $\boldsymbol{V}_{\boldsymbol{\beta}\underline{z}}^{1/2}$ also qualify as the square roots of $\boldsymbol{V}_{\boldsymbol{\alpha}}$ and $\boldsymbol{V}_{\boldsymbol{\beta}} + \boldsymbol{B}\boldsymbol{V}_{\underline{z}}\boldsymbol{B}$. Using one positive and one negative square root matrix allows us to manipulate Equation (A18) to provide the solution;

$$\boldsymbol{\Gamma} = \left(\boldsymbol{I} + \boldsymbol{V}_{\boldsymbol{\beta}\underline{z}}^{1/2} \boldsymbol{V}_{\boldsymbol{\alpha}}^{-1/2}\right)^{-1} \left(\boldsymbol{B} - \boldsymbol{A}\boldsymbol{V}_{\boldsymbol{\beta}\underline{z}}^{1/2} \boldsymbol{V}_{\boldsymbol{\alpha}}^{-1/2}\right) \tag{A22}$$

To confirm that this is a solution, we need to show that it fulfils the second order conditions enshrined in Equations (A19) and (A20). Again inserting one positive and one negative square root matrix into these equations reveals;

$$\left(\boldsymbol{I}+\boldsymbol{V}_{\boldsymbol{\beta}\underline{z}}^{1/2}\boldsymbol{V}_{\boldsymbol{\alpha}}^{-1/2}\right)^{-1}(\boldsymbol{A}+\boldsymbol{B}) \tag{A23}$$

and

$$\left(\boldsymbol{I}+\boldsymbol{V}_{\boldsymbol{\beta}\underline{z}}^{1/2}\boldsymbol{V}_{\boldsymbol{\alpha}}^{-1/2}\right)^{-1}\boldsymbol{V}_{\boldsymbol{\beta}\underline{z}}^{1/2}\boldsymbol{V}_{\boldsymbol{\alpha}}^{-1/2}(\boldsymbol{A}+\boldsymbol{B}) \tag{A24}$$

Our objective is to show that these two matrices are positive definite. Now the matrix $\boldsymbol{V}_{\boldsymbol{\beta}\underline{z}}^{1/2}\boldsymbol{V}_{\boldsymbol{\alpha}}^{-1/2}$ is a diagonal matrix. The elements of the principal diagonal of this matrix consist of the ratio of the positive square roots of the equivalent elements of the diagonals of $\boldsymbol{V}_{\boldsymbol{\alpha}}$ and $\boldsymbol{V}_{\boldsymbol{\beta}}+\boldsymbol{B}\boldsymbol{V}_{\underline{z}}\boldsymbol{B}$. Adding the identity matrix and taking the inverse to form $\left(\boldsymbol{I}+\boldsymbol{V}_{\boldsymbol{\beta}\underline{z}}^{1/2}\boldsymbol{V}_{\boldsymbol{\alpha}}^{-1/2}\right)^{-1}$, does not change the essential properties of this matrix, it is still a diagonal matrix with positive elements along the principal axis. As a result Equations (A23) and (A24) comprise two matrices constructed by pre-multiplying $\boldsymbol{A}+\boldsymbol{B}$ by a positive definite diagonal matrix. Such pre-multiplication simply rescales $\boldsymbol{A}+\boldsymbol{B}$, but does nothing to change its sign definiteness. Since we have assumed that $\boldsymbol{A}$ and $\boldsymbol{B}$ are positive definite matrices and the sum of two positive definite matrices is itself positive definite (Johnson, 1970), we can confirm that both Equation (A23) and Equation (A24) define positive definite matrices. Equation (A22), therefore, is a valid solution consistent with first and second order conditions for an optimum.[62]

In this simple case, the expression describing the linear parameters of the HPF (Equation A21) reduces to;

$$\boldsymbol{\gamma}=\left(\boldsymbol{I}+\boldsymbol{V}_{\boldsymbol{\beta}\underline{z}}^{1/2}\boldsymbol{V}_{\boldsymbol{\alpha}}^{-1/2}\right)^{-1}\left(\overline{\boldsymbol{\alpha}}+\left(\overline{\boldsymbol{\beta}}-\boldsymbol{B}\overline{\underline{z}}\right)\boldsymbol{V}_{\boldsymbol{\beta}\underline{z}}^{1/2}\boldsymbol{V}_{\boldsymbol{\alpha}}^{-1/2}\right) \tag{A25}$$

[62] Note that the solution obtained by taking both positive square roots in the derivation of Equation (A23) is not consistent with the second order conditions defining optimal behaviour.

The simplified model provides some interesting and intuitive insights into the nature of equilibrium in the property market. Imagine a second property market differing only in terms of the strength of preferences for property attributes relative to the numeraire good (x). We could model such a difference by proposing a new mean to the distribution of household preferences, say $\bar{\alpha}'$. If $\bar{\alpha}'$ is larger than $\bar{\alpha}$ then in this second market households place relatively higher weight on property attributes in the determination of their utility. Consequently, from Equation (A25), we see that the equilibrium value for the linear parameters of the HPF, γ, would be relatively more positive. Therefore, in this second market, where property attributes are relatively more highly valued by households, property prices will also be higher for any particular non-zero level of property attributes, z.

In a similar vein, if producing property attributes is relatively more costly for landlords such that $\bar{\beta}' > \bar{\beta}$, then prices will again be higher in the second property market. Finally, if the levels of exogenously supplied property attributes are greater in the second property market such that $\underline{\bar{z}}' > \underline{\bar{z}}$ then property prices will be generally lower.

Further special cases of the general solution in Equations (A18) and (A21) deserve mention:

- *All households are identical*: In this case, there will be no variation in α, the parameter describing variation in preferences across households. As such, α will take the constant value $\bar{\alpha}$ for all households with variance given by $V_\alpha = 0$. From Equation (A17), we see that $V_\alpha = 0$ implies that $\Gamma = -A$, which when replaced in Equation (A21) reveals $\gamma = \bar{\alpha}$. In this special case then, when $\Gamma = -A$ and $\gamma = \bar{\alpha}$ we find that the marginal bid function (Equation A7) for all households is identical and coincides exactly with the equilibrium implicit price function (Equation A13). Since points along the marginal bid function provide the household with constant utility, each of the identical consumers will gain the same utility by choosing to live in a property at any point on the equilibrium hedonic price surface. Heterogeneous landlords will locate at the point on this surface that maximises their profits given their cost function and pair off with any one of the homogenous households.

- *All landlords are identical*: In this case, there is no variation in $\boldsymbol{\beta}$ the parameter describing variation in the cost function of landlords nor any variation in $\underline{z}$, the exogenously provided property attributes. Following a similar logic as previous, we find that $\boldsymbol{\Gamma} = \boldsymbol{B}$ and $\gamma = \overline{\boldsymbol{\beta}} - \boldsymbol{B}\overline{\underline{z}}$, and the marginal offer function (Equation A10) for all landlords will coincide exactly with equilibrium marginal implicit price function (Equation A13). In this case then, identical landlords will earn equal profit wherever they locate on the equilibrium hedonic price surface. Heterogeneous households will choose a location that maximises their utility and pair off with any one of the homogenous landlords.
- *Supply is perfectly inelastic*: In this case, we assume that landlords cannot independently change the attributes of their properties. The distribution of properties in the market is determined solely by the distribution of exogenously supplied property characteristics, $\underline{z}$. In this case, market equilibrium is achieved when the variance of household demand is equal to the variance of exogenously determined property supply. Equation (A17) reduces to;

$$(\boldsymbol{\Gamma} + \boldsymbol{A})^{-1}\boldsymbol{V}_a(\boldsymbol{\Gamma} + \boldsymbol{A})^{-1} = \boldsymbol{V}_{\underline{z}} \tag{A26}$$

denoting the square root of $\boldsymbol{V}_{\underline{z}}$ by the matrix $\boldsymbol{S}_{\underline{z}}$ we obtain the solution;

$$\boldsymbol{\Gamma} = \boldsymbol{S}_a\boldsymbol{S}_{\underline{z}}^{-1} - \boldsymbol{A} \tag{A27}$$

Likewise, equality of the mean of demand and supply requires that Equation (A16) reduces to;

$$(\boldsymbol{\Gamma} + \boldsymbol{A})^{-1}(\overline{\alpha} - \gamma) = \overline{\underline{z}} \tag{A28}$$

which when combined with equation (A27) provides the solution

$$\gamma = \overline{\alpha} - \boldsymbol{S}_a\boldsymbol{S}_{\underline{z}}^{-1}\overline{\underline{z}} \tag{A29}$$

In this case the implicit prices in the property market (Equation A13) take the form;

$$\boldsymbol{p}_z = \overline{\boldsymbol{a}} - \boldsymbol{A}z + \boldsymbol{S}_a \boldsymbol{S}_{\underline{z}}^{-1} (z - \underline{z}) \tag{A30}$$

so that for the "average" property the implicit price function equates to the "average" *inverse demand function*. Moving away from the average property in attribute space, the implicit price function differs from the average inverse demand function by a factor determined by the relative density of preferences for properties at that level of attributes compared to the availability of properties at that level of attributes.

Though the NLQ model abstracts considerably from reality, it provides a closed-form solution that allows us to gain numerous insights to the nature of equilibrium in property markets. In particular it is evident that in all but the simplest cases, equilibrium prices will be a complex function of the distributions of household preferences, landlord costs and exogenously supplied property attributes. In particular, we have seen that the implicit price function does not coincide with the inverse demand function except in the special case in which households are identical in every respect.

A2. More Complex Models of Equilibrium

Whilst the NLQ model is a useful tool for studying equilibrium it is both too restrictive in its imposition of functional forms and too liberal in terms of the economic behaviour that it allows. In a series of recent papers, Ivar Ekeland, James Heckman, Rosa Matzkin and Lars Nesheim (Ekeland et al., 2003; Heckman et al., 2002; Heckman et al., 2003) have investigated the impacts on the equilibrium price function of allowing for more flexible functional forms and imposing economically plausible behaviour. Since, these problems no longer provide closed-form solutions, numerical methods are used to approximate the HPF; the solution to the differential equation in Equation (29) in Chapter 1. Their research shows that even minor perturbations from the assumptions of the NLQ model disrupt the simple quadratic form of the equilibrium HPF.

In Ekeland et al. (2003) the normality of the distribution of the preference and cost parameters (α and β) is relaxed by modelling these heterogeneity components as the mixture of two normal distributions. Rather than the constant curvature found with the quadratic HPF of the NLQ model, models in which the heterogeneity is non-normal produce price functions with far from constant implicit prices. Indeed, the further the distribution of heterogeneity is from normal the further the curvature strays from constancy. Imposing economically reasonable restrictions (i.e. positive implicit prices, only positive quantities of attributes demanded and supplied) only serves to exaggerate the nonlinearity of implicit prices. Indeed as Ekeland et al. (2003) prove in the context of the NLQ model, the equilibrium HPF is generically nonlinear.

Heckman et al. (2003) also examine the density of properties with different levels of attributes supplied to the market. With the NLQ model, this density follows a normal distribution, as the heterogeneity components are made less and less normal, the density follows suit. In all cases illustrated, however, the density of supply of properties with different levels of attributes remains a fairly regular unimodal distribution.

Heckman et al. (2003) extend this investigation by examining models in which the linear and quadratic terms in the utility and cost functions (Equations A1 and A8, respectively) are replaced with higher order polynomials. This allows far greater flexibility in the shapes of these two functions. Again, the equilibrium HPF is shown to be highly nonlinear. Moreover, the density of supply in the illustrated cases is far from normal exhibiting many modes. As Heckman et al (2003) point out, "the model is capable of generating equilibria in which there are nearly gaps in the range of products marketed".

The upshot of this research is to highlight the fact that real world equilibrium HPFs are likely to be highly nonlinear. Moreover, it is to be expected that in equilibrium the supply of properties boasting different levels of attributes may well be extremely lumpy.

A3. Interaction-Based Models of Equilibrium

A fundamental assumption of the models described in the previous sections is that the value that households place on properties is determined by the characteristics of those properties and the locations in which they are set. A number of authors (including Epple and Platt, 1998; Epple and Seig, 1999a, 1999b, 2002; Nesheim, 2002) have investigated the nature of equilibrium in the property market when households base their valuation of properties not on the characteristics of the locations themselves but on the characteristics of the equilibrium sets of people in those locations. Here we review the model of Nesheim (2002) to explore the implications of such interaction-based models.

As in the previous models, the property market is assumed to consist of profit-maximising landlords and utility-maximising households. Unlike the previous models, however, Nesheim's model is one of locational choice. To make things simple it is assumed that location can be treated as one dimensional, such that locations in the property market can be indexed by $l \in R_+$. Each location boasts a set of properties owned by competitive landlords. Households choose between these locations based on the set of people that also chose to live at that location. We desire to characterise the HPF, $\widetilde{P}(l)$, that brings the market into equilibrium.

On the supply side of the market we assume that the quantity of properties available at each location, *l*, is fixed. To focus on the interaction effects between households, we ignore differences in the characteristics of the properties (and locations) themselves. Thus these characteristics are assumed not to influence the utility of households or are assumed to be identical for all properties. In the model, the supply of properties is described by a positive continuous density function $h(l)$.[63]

Our focus is on the demand side of the market. It is assumed that households are prepared to pay more to live in locations with higher levels of school quality since

---

[63] Since we will subsequently be describing the households in the property market as a multidimensional probability density function (in which the different dimensions represent different household characteristics), the assumption that $\int h(l)dl \geq 1$ is required to ensure that there are sufficient properties to house the population.

this is an input into the educational attainment of their children. For the purposes of the model, it is assumed that school quality is determined by the average educational attainment of parents in a location. Thus, the higher the average schooling of parents choosing to live in a location, the higher the quality of schools and the more desirous households are to reside in that location. Let us introduce the following notation;

- s_0 is the schooling attainment of the *parents* in a household
- s_1 is the schooling attainment of the *children* in a household
- $S(l)$ is the average (parental) schooling of the households choosing to live at location *l*.

Now, the schooling attainment of children is assumed to be produced through the simple Cobb-Douglas technology;

$$s_1(l) = a_1 S(l)^{\eta_1} s_0^{\eta_2} \tag{A31}$$

where a_1 is the child's ability and η_1 and η_2 are the parameters of the schooling attainment function.[64]

In the simplest model, households are assumed to choose a location, *l*, and an amount to spend on other goods, *x*, so as to maximise the quasilinear utility function;

$$\begin{aligned} U(l,x;\boldsymbol{\alpha}) &= \kappa s_1(l) + x \\ &= \kappa s_1(l) + y - \widetilde{P}(l) \\ &= \kappa a_1 S(l)^{\eta_1} s_0^{\eta_2} + y - \widetilde{P}(l) \end{aligned} \tag{A32}$$

As in the NLQ model, *y* is household income and $\boldsymbol{\alpha}$ is the vector describing the particular tastes and characteristics of the household. We shall define this vector more exactly shortly. Here we note that one of the elements of $\boldsymbol{\alpha}$ when transformed, gives κ a parameter measuring the household's strength of preferences for their children's educational attainment relative to their own consumption.

---

[64] For reasons that are not required for our exposition, Nesheim (2002) includes a further element in the schooling attainment function representing a random shock.

The utility function in Equation (A32) describes households according to four traits; income (y), parental education (s_0), children's ability (a_1) and preferences for children's educational attainment (κ). Let us assume that these characteristics are distributed amongst households in the property market according to a multivariate lognormal density function. Nesheim defines the vector $\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \end{bmatrix}'$ where $\alpha_1 = \ln y$, $\alpha_2 = \ln s_0$, $\alpha_3 = \ln a_1$ and $\alpha_4 = \ln \kappa$. Therefore, the characteristics of each household in the property market can be summarised in a vector $\boldsymbol{\alpha}$. Since the traits themselves are distributed lognormally, we can describe the distribution of households in the property market by;

- the distribution of the vector $\boldsymbol{\alpha}$ which is multivariate normal with mean $\bar{\boldsymbol{\alpha}}$ and variance $\boldsymbol{V_\alpha}$

An important part of the model is the *equilibrium assignment function*, $\widetilde{G}(\boldsymbol{\alpha})$, and its inverse, the relation $\widetilde{G}^{-1}(l)$. We define these as follows;

- $l = \widetilde{G}(\boldsymbol{\alpha})$ is the function describing the optimal choice of location, l, for a household with characteristics $\boldsymbol{\alpha}$, when the market is in equilibrium.

- $\widetilde{G}^{-1}(l)$ defines the characteristics of the set of households that in equilibrium choose to locate at l according to $\left\{\boldsymbol{\alpha}(l) \mid \boldsymbol{\alpha} \in \widetilde{G}^{-1}(l)\right\}$.

The equilibrium assignment function allows us to define the quality of schools in any particular location. Since, $\widetilde{G}^{-1}(l)$ informs us which households will choose to locate at l, the average parental schooling at l is given by;

$$\begin{aligned} S(l) &= E\left[e^{\alpha_2} \mid \boldsymbol{\alpha} \in \widetilde{G}^{-1}(l)\right] \\ &= E\left[s_0 \mid \boldsymbol{\alpha} \in \widetilde{G}^{-1}(l)\right] \end{aligned} \tag{A33}$$

As a result the equilibrium assignment function enters the households utility function by defining the quality of schools in each location when the market is in equilibrium.

All the essential elements of the model are now in place. Households set out to choose a location that maximises their utility as defined by Equation (A32). The solution to this maximisation problem is a demand correspondence;

$$\{l\} = l^d\left(\alpha, \widetilde{G}, \widetilde{P}\right) \tag{A34}$$

where $\{l\}$ is the set of locations that maximise the utility of a household with characteristics α, given the equilibrium assignment function, $\widetilde{G}(\alpha)$, and equilibrium price function, $\widetilde{P}(l)$.

An equilibrium in this model, is established by two relations, $\widetilde{G}^{-1}(l)$ and $\widetilde{P}(l)$, which ensure that;

Condition A: if $\widetilde{G}^{-1}(l)$ indicates that households with characteristics α will choose to live at location *l*, then *l* must also be an optimising choice of location for those same households, given the price function $\widetilde{P}(l)$. More formally;

$$\text{if} \quad \alpha \in \widetilde{G}^{-1}(l) \quad \text{then} \quad l \in l^d\left(\alpha, \widetilde{G}, \widetilde{P}\right) \quad \text{for all } \alpha$$

If Condition A is met then no household has any incentive to move to a property at another location. As such the property market is in a state of equilibrium.

Moreover, an equilibrium must meet three further conditions

Condition B: the density of demand for properties at all location must not exceed the density of supply of properties.

Condition C: demand must equal supply at all locations with a positive price.

Condition D: landlords profits must not be negative.

Nesheim (2002) shows that an intuitive feature of an equilibrium fulfilling these conditions is that;

- If households with the same characteristics, α, choose to live in two different locations, then these two locations must be of the same quality (i.e. have the same average parental schooling, *S*).

- If two locations have the same quality then they must also command the same price.

Importantly, these observations allow Nesheim to restate the optimisation problem in terms of choice of locational quality (*S*) as opposed to one of location itself (*l*). In particular, imagine a locational equilibrium defined by two functions $\widetilde{G}'(l)$ and $\widetilde{P}'(l)$. It is quite possible that a second locational equilibrium exists defined by the two functions $\widetilde{G}''(l)$ and $\widetilde{P}''(l)$ in which households choose to live in different locations but choose the same quality of location and pay the same price to live at that location. Whilst the second locational equilibrium might change the price and quality at a particular location, *l*, it has no impact on the outcome for households. Whilst each household enjoys the same utility in the two equilibria the same is not true for landlords. In effect, the two equilibria imply alternate distributions of wealth across landlords of properties in different locations.

From the point of view of the household, our object of study, these separate *locational equilibria* belong to the same *quality sorting equilibrium*, in that they maintain the same quality-price schedule and allocate the same households to each neighbourhood quality level.

We can define a quality sorting equilibrium in much the same way we did a locational equilibrium. The equilibrium is established by two relations that mimic their locational counterparts, $\widetilde{G}^{-1}(l)$ and $\widetilde{P}(l)$;

- $G^{-1}(S)$ which defines the characteristics of the set of households choosing a quality level, and its inverse, $G(\boldsymbol{a})$ the *quality assignment function* which maps households to quality levels
- $P(S)$ the *quality price function* determining the price paid for a property in a location of a particular quality.

A quality sorting equilibrium arises when the following conditions are met;

Condition 1: For every quality level, *S*, the function $G^{-1}(S)$ assigns a set of households to a location whose average schooling equals *S* itself. More formally;

$$E\left[e^{\alpha_2} \mid \boldsymbol{\alpha} \in G^{-1}(S)\right] = S \quad \text{for all levels of } S$$

***Condition** 2*: For each household, $\boldsymbol{\alpha}$, the function $G(\boldsymbol{\alpha})$ must indicate a chosen level of locational quality that maximises the households utility given the price function $P(S)$. More formally;

$$\text{if } \ G(\boldsymbol{\alpha}) = S \ \text{ then } \ S \in \underset{S' \in G(\boldsymbol{\alpha})}{\arg\max}\, U(y - P(S'), S', \boldsymbol{\alpha}) \quad \text{for all households } \boldsymbol{\alpha}$$

Nesheim proves that for every *quality sorting equilibrium* satisfying conditions 1 and 2, there exists a *locational equilibrium* satisfying conditions *A*, *B*, *C* and *D*.

The household's problem can now be restated as maximising the utility function;

$$U(S;\boldsymbol{\alpha}) = \kappa a_1 S^{\eta_1} s_0^{\eta_2} + y - P(S) \tag{A35}$$

with respect to the choice of locational quality *S*. We obtain the first order condition:

$$\begin{aligned} p_S &= \eta_1 S^{\eta_1 - 1} s_0^{\eta_2} a_1 \kappa \\ &= \eta_1 S^{\eta_1 - 1} e^{\eta_2 \alpha_2 + \alpha_3 + \alpha_4} \end{aligned} \tag{A36}$$

where p_S is the derivative of the price function with respect to locational quality (the implicit price of locational quality). Again we assume that the second order conditions are globally satisfied and will check the veracity of this assumption once we have solved for the equilibrium price function.

Equation (A36) presents the standard conditions for optimal choice; households choose a level of locational quality that equates marginal cost with marginal benefit. Here cost is the increased price of property and benefit is the increased educational attainment of children. Given the normalisation of the utility function, both costs and benefits are measured in money.

Taking the natural logarithm of Equation (A36) recasts the first order conditions as;

$$\begin{aligned} \ln p_S - \ln \eta_1 + (1-\eta_1)\ln S &= \eta_2\alpha_2 + \alpha_3 + \alpha_4 \\ f(S) &= \boldsymbol{A'\alpha} \end{aligned} \tag{A37}$$

where $\boldsymbol{A} = [0 \quad \eta_2 \quad 1 \quad 1]$. The rearrangement of terms in Equation (A37) allows us to characterise optimal choice as the matching of two indices;

- $f(S)$ isolates all those elements of the FOCs that are dependent on locational quality, *S*, and are constant across all households. Roughly speaking it is an index summarising the marginal costs of increasing locational quality net of the benefits enjoyed from that increase enjoyed by all households independent of their characteristics.

- $\boldsymbol{A'\alpha}$ isolates all those elements of the FOCs that are specific to a particular household. Roughly speaking, it is an index describing the marginal benefits of increasing locational quality that are specific to a household. Since $\boldsymbol{A'\alpha}$ is the log of the monetary terms describing the marginal benefits of locational quality to the household, Nesheim (2002) characterises $\boldsymbol{A'\alpha}$ as an index describing households' marginal willingness to pay (WTP).[65]

The FOCs as summarised in Equation (A37), allow us to characterise $G^{-1}(S)$, the first of the two relations defining a market equilibrium. In particular note that all households choosing a particular level of locational quality $\hat{S}$ must be those for whom the FOC $f(\hat{S}) = \boldsymbol{A'\alpha}$ holds true. Of course, numerous values of the vector $\boldsymbol{\alpha}$ could meet this condition. So that in equilibrium $G^{-1}(S)$ is the set defined by the following relation;

$$G^{-1}(S) = \{\boldsymbol{\alpha} \mid f(S) = \boldsymbol{A'\alpha}\} \tag{A38}$$

Given a level of locational quality *S*, and an equilibrium price function $P(S)$, $G^{-1}(S)$ describes the characteristics of the set of households choosing that level of

---

[65] Marginal WTP for locational quality is actually given by the expression $\eta_1 \, S^{\eta_1 - 1} \, e^{\boldsymbol{A'\alpha}}$.

locational quality. Since this definition of $G^{-1}(S)$ is derived directly from the FOCs we can be sure that we have satisfied ***Condition 2*** of a quality sorting equilibrium.

All that remains is to satisfy ***Condition 1*** and to solve for the equilibrium price function. *Condition 1* demands that the function $G^{-1}(S)$ assigns a set of households to a location whose average schooling equals *S* itself. That is;

$$S = E\left[e^{\alpha_2} \mid \boldsymbol{\alpha} \in G^{-1}(S)\right] \tag{A39}$$

Armed with a relation describing the equilibrium set of households at each locational quality level (Equation A38), we are now able to provide an explicit representation of the RHS of Equation (A39). To do so requires establishing the distribution of the log of parental schooling ($\alpha_2 = \ln s_0$) at each level of locational quality.

We know from Equation (A38) that the set of households choosing a particular level of locational quality, *S*, must be those for which $f(S) = \boldsymbol{A'\alpha}$. Since we have assumed that the vector $\boldsymbol{\alpha}$ is distributed as a multivariate normal distribution with mean $\bar{\boldsymbol{\alpha}}$ and variance $\boldsymbol{V_\alpha}$, $\boldsymbol{A'\alpha}$ is simply a linear combination of normal random variates. As a result the conditional distribution of α_2, the second element $\boldsymbol{\alpha}$, given $f(S) = \boldsymbol{A'\alpha}$ is also distributed normally according to;

$$\left(\alpha_2 \mid f(S) = \boldsymbol{A'\alpha}\right) \sim N\left(\mu_S, \sigma_S^2\right) \tag{A40}$$

where;

$$\mu_S = \bar{\alpha}_2 + \left(\boldsymbol{A'V_\alpha e_1}\right)\left(\boldsymbol{A'V_\alpha A}\right)^{-1}\left(f(S) - \boldsymbol{A'\bar{\alpha}}\right) \tag{A41}$$

and

$$\sigma_S^2 = V_{\alpha_2} - \left(\boldsymbol{A'V_\alpha e_1}\right)^2\left(\boldsymbol{A'V_\alpha A}\right)^{-1} \tag{A42}$$

Here μ_S and σ_S^2 are the conditional mean and variance of α_2 at any particular level of locational quality, $\bar{\alpha}_2$ and V_{α_2} are the population mean and variance of α_2, and $\boldsymbol{e}_1$ is the vector $\begin{bmatrix}0 & 1 & 0 & 0\end{bmatrix}'$.

Since α_2 is distributed normally, $s_0 = e^{\alpha_2}$ is distributed log normally and we arrive at an explicit representation for Equation (A39);

$$S = E\left[e^{\alpha_2} \mid \boldsymbol{\alpha} \in G^{-1}(S)\right] = \exp\left(\mu_S + \sigma_S^2\right) \tag{A43}$$

Therefore, the quality of each neighbourhood as defined by the average level of parental schooling attainment is a function of $\bar{\boldsymbol{\alpha}}$ and $\boldsymbol{V}_{\boldsymbol{\alpha}}$, the mean and variance of household traits in the population, η_1 and η_2, the parameters of the schooling attainment function, as well as locational quality, *S*, and prices, p_S, through $f(S)$.

Replacing Equation (A41) and (A42) in Equation (A43) and solving for p_S provides the expression;

$$p_S = S^{\left(\frac{1}{L_1}+\eta_1-1\right)} L_0^{-\frac{1}{L_1}} \tag{A44}$$

where $L_0 = \exp\left(\bar{\alpha}_2 - L_1\left(\ln\eta_1 + A\bar{\boldsymbol{\alpha}}\right) + V_{\alpha_2}/2\right)$ and $L_1 = \left(A'\boldsymbol{V}_{\boldsymbol{\alpha}}\boldsymbol{e}_1\right)\left(A'\boldsymbol{V}_{\boldsymbol{\alpha}}A\right)^{-1}$ which are two constants depending on the parameters of the model and the distributions of household traits. Notice that L_1 is simply the proportion of the total variance in the index of WTP for neighbourhood quality, $\boldsymbol{A}'\boldsymbol{\alpha}$, that is accounted for by the variance in the log of parental schooling, $\ln s_0 = \alpha_2$. Put simply, L_1 measures the correlation between parental schooling attainment and WTP.

Finally, we can establish the equilibrium price function that underlies Equation (A43) and thereby conforms to ***Condition 1*** of our conditions for an equilibrium;

$$
\begin{aligned}
P(S) &= \int p_S \, dS \\
&= L_0^{-\frac{1}{L_1}} \int S^{\left(\frac{1}{L_1}+\eta_1-1\right)} dS \qquad \text{(A45)} \\
&= L_0^{-\frac{1}{L_1}} \left(\eta_1 + \frac{1}{L_1} \right)^{-1} S^{\eta_1+\frac{1}{L_1}}
\end{aligned}
$$

To confirm this price function supports an optimum solution to the households problem, replace the implicit price function (Equation A44) in the FOC (Equation A36) and differentiate with respect to the choice of locational quality to form the second order condition;

$$
\begin{aligned}
&(\eta_1 - 1)\eta_1 S^{\eta_1-2} e^{A'\alpha} - \left(\eta_1 - 1 + \frac{1}{L_1} \right) S^{\frac{1}{L_1}+\eta_2-2} L_0^{-\frac{1}{L_1}} < 0 \\
&= S^{-1} \left((\eta_1 - 1) p_S - \left(\eta_1 - 1 + \frac{1}{L_1} \right) u_s \right) < 0
\end{aligned}
\qquad \text{(A46)}
$$

where u_S is simply the marginal benefits of increased locational quality as defined by the RHS of Equation (A36). Since the FOC (Equation A36) stipulates that $p_S = u_S$, it is easily established that the SOC in Equation (A46) is always satisfied so long as $\eta_1 > 0$ and $L_1 > 0$. That is the conditions for an optimum are achieved so long as two perfectly intuitive conditions are fulfilled; on average parents with more education must be WTP more for school quality and the marginal produce of school quality (i.e. the increase in the schooling attainment of children, s_0, from increases in locational quality, S) must be positive.

What does Nesheim's model reveal concerning the nature of a quality sorting equilibrium? First, observe the equilibrium price function in Equation (A45). The elasticity of prices with respect to locational quality is given by the expression $\eta_1 + L_1^{-1}$. If $\eta_1 + L_1^{-1} < 1$ then the price function is concave whilst if $\eta_1 + L_1^{-1} \geq 1$ the price function will be convex.

Recall from the production function of Equation (A31) that η_1 is the elasticity of children's schooling attainment with respect to school quality whilst L_1 is the correlation between household WTP and parents own schooling attainment. Clearly larger values of η_1 will precipitate larger price differences across neighbourhoods of increasing quality. The greater the importance of school quality in determining children's schooling attainment, the greater the price differentials required to ensure households segregate into locations of their preferred quality. In contrast larger values of L_1 result in smaller price differences across neighbourhoods of increasing quality. The greater the correlation between WTP and parents schooling attainment, the more homogenous are households seeking to locate in neighbourhoods of the same quality. As a result smaller price differentials are required to maintain the equilibrium segregation of households.

A further interesting result can be found by examining the distribution of household types at a particular level of locational quality. Let us begin by examining the conditional distribution of log parental educational attainment, $\ln s_0 = \alpha_2$, as described in Equation (A40). Armed with an explicit expression for p_S (Equation A44) the mean and variance of this distribution (Equations A41 and A42) simplify to;

$$\mu_S = \ln S - \sigma_S^2 / 2 \tag{A47}$$

and

$$\sigma_S^2 = V_{\alpha_2}\left(1 - \rho_{\alpha_2}^2\right) \tag{A48}$$

where ρ_{α_2} is the correlation between log parental educational attainment and WTP for neighbourhood quality.[66]

σ_S^2, the conditional variance of log parental schooling is constant across neighbourhoods. Notice, however, that the variance within a neighbourhood is smaller than the population variance, V_{α_2}, by a factor $1-\rho_{\alpha_2}^2$. As Nesheim states, "Since people sort based on common WTP and since that WTP is correlated with parental education, individual neighbourhoods are more homogenous in terms of education than the population at large" (Nesheim, 2002; p.31). The degree of within-

[66] This differs from L_1 only by a factor of η_1.

neighbourhood homogeneity is determined by ρ_{α_2}. The more highly WTP is correlated with parental education, the smaller the variance of this trait within neighbourhoods. In the limit, when WTP is almost perfectly correlated with parental education, all households in a neighbourhood will have the same level of parental education. In this case sorting based on WTP is equivalent to sorting based on parental education.

A similar analysis is possible for log children's ability, $\alpha_3 = \ln a_1$. Here the conditional distribution is;

$$\left(\alpha_3 | f(S) = A'\boldsymbol{\alpha}\right) \sim N\left(\mu_a, \sigma_a^2\right) \tag{A49}$$

where;

$$\mu_a = \overline{\alpha}_3 + \frac{\sqrt{V_{\alpha_3}}\rho_{\alpha_3}}{\sqrt{V_{\alpha_2}}\rho_{\alpha_2}}\left(\ln S - \overline{\alpha}_2 - \sigma_S^2 / 2\right) \tag{A50}$$

and

$$\sigma_a^2 = V_{\alpha_3}\left(1 - \rho_{\alpha_3}^2\right) \tag{A51}$$

Here V_{α_3} is the population variance of log-ability and ρ_{α_3} is the correlation between log-ability and WTP for neighbourhood quality.

Once again, the greater the degree of correlation between log-ability and WTP the more homogenous households are within a community with regards to the ability of their children.

Also from Equation (A50) we can observe how the average log-ability in each neighbourhood is affected by the parameters of the model. Since, satisfaction of the SOCs demands that $\rho_{\alpha_2} > 0$ we can be sure that average log-ability is increasing with neighbourhood quality so long as $\rho_{\alpha_3} > 0$. That is, so long as WTP is positively correlated with log-ability, neighbourhoods of higher locational quality will be characterised by having higher average ability amongst children. Likewise, the greater the population variance of log-ability, V_{α_3}, the more rapidly average log-ability increases with neighbourhood quality. This observation reflects the fact that holding other things constant, a larger variance of log-ability must increase the correlation between log-ability and WTP since such an increase would make log-ability a larger component of the variance of WTP. In short, the more closely related log-ability is to WTP for locational quality, the greater will be the differences in average abilities across neighbourhoods.

Nesheim's model of locational choice gives us numerous insights into the nature of equilibrium in a property market where households value the sets of people that choose to reside in their neighbourhood. In particular, the model predicts that such behaviour will result in a sorting equilibrium in which the traits of households in a neighbourhood are likely to be less varied than those of the population as a whole. Indeed, the more correlated a trait is with WTP for locational quality, the more homogenous neighbourhoods are likely to be in this trait and the greater will be the differences in the average level of this trait across neighbourhoods.

In the same paper Nesheim (2002) also investigates a generalisation of the model in which the simple linear utility function of Equation (A35) is replaced by an exponential utility function. As with the NLQ model such a generalisation results in a problem that does not possess a closed-form solution. Consequently, Nesheim uses numerical techniques to characterise equilibrium. Equilibrium in the more general model displays much richer patterns of sorting and much more complex shapes of equilibrium price function. Indeed, Nesheim shows that for some values of the parameters he must allow for a kinked price function. As with the NLQ model, Nesheim shows that the more complex model can result in somewhat lumpy provision of locational quality. In Nesheim's model relatively few households choose to live in neighbourhoods of moderate locational quality. The majority of households choose relatively low or relatively high levels of locational quality and these populations are characterised by markedly different average traits and a markedly different slope to the equilibrium HPF.

Appendix B: Factor Scores

Table B1: Factor Score Coefficients

Attribute	Factor 1 Wealth	Factor 2 Ethnicity	Factor 3 Age	Factor 4 Family
No car	-0.45246	-0.05167	0.07424	-0.04367
Two cars	0.09231	-0.04601	-0.04466	0.01243
Unemployment	-0.07536	0.01154	0.02635	-0.01158
Non-owners	-0.14204	-0.09554	0.03441	0.08386
One-parent families	-0.04458	-0.08744	-0.0356	-0.02097
Low Social Class	-0.01139	0.00756	0.00409	-0.01719
Families	0.02263	0.02318	0.01858	-0.05969
Age 0 to 10	-0.01938	-0.19307	-0.19924	0.93989
Age 11 to 17	0.10494	0.09869	0.06699	0.52499
Age 18 to 24	0.06287	0.16916	-0.47298	-0.21339
Age 25 to 34	0.13391	0.01061	-0.67956	-0.15515
Age 35 to 49	0.29335	-0.06849	-0.30182	0.1407
Age 50 to 64	0.08283	0.1746	0.04727	-0.15016
Age ≥ 65	0	0	0	0
Over Crowding	0.0152	0.09753	0.02219	-0.09834
Non White	-0.04687	0.9091	-0.23225	-0.45231
Black	-0.04451	-0.12749	-0.02446	0.06222
Asian	0.17711	0.04704	0.22543	0.04086

APPENDIX C: PLSC & PLSS PARAMETER ESTIMATES

Table C1: Parameter Estimates for the Spatial Constants Partial Linear Model

Variable	Parameter Estimates by Submarket (p-values in brackets)							
	1	2	3	4	5	6	7	8
Nonparametric Variables:								
ln(Floor Area)	0.3301	0.4872	0.3752	0.3518	0.4638	0.4803	0.5878	0.6024
ln(Garden Area)	0.0668	0.1124	0.1036	0.1453	0.2335	0.2743	0.1347	0.1802
Age	-0.0141	-0.0053	0.0019	-0.0006	-0.0387	-0.0084	-0.0057	-0.0274
Sale Date	0.0001	0.0002	0.0002	0.0001	0.0001	0.0001	0.0002	0.0002
Structural Variables:								
Bedrooms 1	0.0558 (.747)	-0.0851 (.667)	0.0389 (.735)	0.0547 (.427)	-0.2393 (.235)		-0.0792 (.646)	-0.4746 (.000)
Bedrooms 2	0.0169 (.362)	0.0037 (.842)	-0.0252 (.046)	-0.0082 (.644)	-0.0304 (.289)	-0.0007 (.992)	-0.0319 (.002)	-0.0444 (.054)
Bedrooms 3	base case	base case	base case	base case	base case	base case	base case	base case
Bedrooms 4	-0.0347 (.609)	0.0658 (.163)	0.0341 (.331)	0.0473 (.074)	0.0447 (.103)	0.0279 (.468)	0.0184 (.574)	0.0465 (.000)
Bedrooms 5		0.0568 (.642)	0.2204 (.015)		0.0948 (.061)	0.0189 (.682)	0.2323 (.017)	0.0648 (.014)
Bedrooms 6		-0.0796 (.748)	0.1717 (.389)		-0.0308 (.703)	0.0344 (.558)		0.2553 (.007)
Bedrooms 7					-0.0614 (.788)	-0.01 (.912)		0.245 (.079)
Bedrooms 8						-0.0018 (.986)		
Bedrooms 9						-0.1721 (.496)		
Bedrooms 10						-0.6208 (.062)		
Bedrooms 11						-0.257 (.344)		
Bedrooms 12						0.3536 (.458)		
WCs 1	base case	base case	base case	base case	base case	base case	base case	base case
WCs 2	0.0302 (.112)	0.026 (.195)	0.04 (.004)	0.01 (.482)	0.0154 (.384)	-0.0246 (.382)	-0.0138 (.144)	-0.0104 (.324)

Variable	Parameter Estimates by Submarket (p-values in brackets)							
	1	2	3	4	5	6	7	8
WCs 3	0.0793 (.645)		0.1081 (.253)	0.0126 (.836)	-0.0074 (.963)	0.0715 (.246)	0.2075 (.027)	0.0219 (.449)
WCs 4				-0.0578 (.753)		0.0313 (.822)		
WCs 5						0.0033 (.989)		
Floors 2	base case	base case	base case	base case	base case	base case	base case	base case
Floors 3	-0.0271 (.439)	-0.0283 (.786)	0.0009 (.982)	-0.0874 (.012)	-0.0729 (.130)	-0.1974 (.001)	-0.0225 (.405)	-0.2703 (.000)
Floors 4			-0.1744 (.128)		-0.1627 (.055)	-0.6445 (.000)		-0.5308 (.000)
Floors 5			0.1528 (.579)			-0.9817 (.000)		
Floors 6						-0.3626 (.102)		
Floors 7						26.12 (.895)		
Garage	0.0549 (.000)	-0.0313 (.399)	0.0639 (.001)	0.0674 (.000)	0.0676 (.000)	0.0116 (.764)	0.031 (.000)	0.0597 (.000)
Central Heating	0.0936 (.000)	0.0066 (.805)	0.1499 (.000)	0.0822 (.002)	0.0642 (.090)	0.0307 (.780)	0.0745 (.001)	0.2059 (.000)
Detached Bungalow	0.3827 (.099)			0.0785 (.071)	0.0521 (.548)	0.1253 (.124)	0.0985 (.030)	0.1585 (.000)
Semi-Detached Bungalow	-0.2974 (.004)			0.0338 (.321)	-0.0507 (.769)		0.133 (.016)	0.033 (.429)
End Terrace Bungalow	-0.0515 (.767)		-0.2209 (.183)	-0.2224 (.277)			0.1191 (.521)	
Terrace Bungalow							0.0005 (.997)	0.1849 (.412)
Detached House	0.0063 (.897)	-0.0366 (.811)	0.1209 (.008)	0.1513 (.000)	0.0737 (.010)	0.1664 (.000)	0.1218 (.000)	0.1049 (.000)
Semi-Detached House	base case	base case	base case	base case	base case	base case	base case	base case
End Terrace House	-0.0543 (.000)	-0.1137 (.035)	-0.0277 (.271)	-0.0928 (.000)	0.0108 (.696)	-0.2467 (.008)	-0.0777 (.000)	-0.073 (.008)
Terrace House	-0.0857 (.000)	-0.1365 (.005)	-0.051 (.025)	-0.0712 (.000)	-0.0051 (.826)	0.0319 (.602)	-0.0793 (.000)	-0.0578 (.022)
BG1			0.0201 (.921)	0.0461 (.807)			-0.094 (.528)	-0.2937 (.080)
BG2				0.1923 (.313)		-0.0286 (.818)		0.4602 (.006)
BG3	-0.012 (.804)	-0.0389 (.128)	-0.019 (.436)	-0.0069 (.932)	0.1551 (.031)		0.0756 (.127)	0.0371 (.827)

Variable	Parameter Estimates by Submarket (p-values in brackets)							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
BG4	0.0147 (.639)	base case	base case	0.0936 (.171)	base case	0.0971 (.265)	0.0141 (.583)	0.0865 (.012)
BG5			-0.2305 (.052)		0.0901 (.056)	0.0431 (.667)		0.2647 (.001)
BG8	0.0625 (.483)	0.1388 (.075)	0.0258 (.301)	-0.059 (.556)	0.0728 (.078)	-0.3079 (.020)	0.0652 (.308)	0.103 (.019)
BG9		-0.1788 (.460)			0.1536 (.039)	-0.0288 (.708)		0.1436 (.004)
BG10						-0.1054 (.211)		0.2223 (.052)
BG19	-0.8651 (.000)		0.0467 (.654)	0.1768 (.063)	0.0824 (.187)	0.0514 (.583)	0.1376 (.023)	0.0833 (.004)
BG20		0.1916 (.013)	0.0185 (.669)	-0.1181 (.051)	0.123 (.005)	-0.2283 (.057)	-0.0687 (.000)	-0.0028 (.931)
BG21	0.0882 (.000)	0.2016 (.005)	0.0506 (.219)	0.0111 (.849)	0.1485 (.001)	-0.1136 (.042)	base case	base case
BG24			-0.3316 (.097)		0.2163 (.012)	base case	0.0911 (.207)	0.1458 (.000)
BG25						0.1659 (.233)		
BG30	0.0233 (.443)	0.1301 (.285)	-0.0519 (.710)	-0.1426 (.008)	0.1576 (.027)	-0.321 (.015)	-0.0338 (.117)	-0.0587 (.093)
BG31	0.069 (.207)	0.957 (.241)	0.3822 (.003)	base case	0.2521 (.053)	-0.0317 (.812)	-0.0127 (.812)	-0.1652 (.010)
BG32	0.0494 (.795)		0.6146 (.000)	0.0627 (.135)	0.5175 (.000)	0.0306 (.748)	-0.0361 (.497)	-0.0864 (.173)
BG35						0.184 (.232)		-0.0048 (.960)
BG36	-0.0342 (.711)	0.8219 (.336)	0.2133 (.160)	-0.1063 (.216)	-0.2493 (.112)			-0.1671 (.359)
Neigbourhood Variables:								
Wealth	-0.1208 (.000)	-0.0046 (.883)	-0.047 (.010)	-0.1042 (.000)	-0.1022 (.000)	-0.1696 (.000)	-0.0546 (.001)	-0.0896 (.000)
Ethnicity	-0.0945 (.002)	0.039 (.055)	-0.0433 (.028)	-0.0904 (.020)	-0.0311 (.025)	-0.2084 (.000)	-0.0799 (.002)	-0.106 (.002)
Age	0.0017 (.914)	0.0384 (.121)	0.0014 (.899)	0.0139 (.189)	0.0372 (.025)	0.0673 (.000)	0.0162 (.041)	0.0592 (.000)
Family	-0.0316 (.005)	-0.0315 (.054)	-0.0179 (.090)	-0.0466 (.000)	-0.0261 (.053)	-0.0313 (.102)	-0.0112 (.291)	-0.0255 (.064)
Environmental Variables:								
Road Noise	0.002 (.134)	0.0019 (.322)	-0.005 (.000)	-0.0029 (.033)	-0.0061 (.000)	-0.0036 (.156)	-0.0031 (.000)	-0.0032 (.006)

Variable	Parameter Estimates by Submarket (p-values in brackets)							
	1	2	3	4	5	6	7	8
Rail Noise	-0.0097 (.045)	-0.0035 (.450)	-0.0091 (.006)	-0.0124 (.001)	-0.0139 (.007)	-0.0128 (.248)	-0.0005 (.879)	-0.0078 (.037)
Air Noise	-0.0139 (.009)		0.0089 (.655)	0.0032 (.440)	0.0406 (.034)	0.0311 (.668)	-0.0062 (.305)	0.0033 (.846)
Park Views	-0.0001 (.499)	0.0005 (.159)	-0.0003 (.120)	0.0002 (.312)	9E-06 (.964)	-0.0001 (.608)	-2E-05 (.845)	0.0001 (.145)
Water Views	0.0063 (.004)	-0.0138 (.163)	0.0005 (.269)	-0.0103 (.002)	0.0046 (.141)	0.0017 (.373)	-0.0025 (.342)	4E-06 (.992)
Locational Variables:								
CBD	-7E-07 (.988)	0.0003 (.009)	-0.0001 (.035)	-2E-05 (.591)	-0.0002 (.033)	0.0002 (.127)	1E-05 (.673)	-6E-05 (.104)
Airport	-4E-06 (.934)	-0.0003 (.002)	0.0001 (.124)	-9E-05 (.018)	-1E-06 (.987)	-0.0003 (.023)	-8E-05 (.006)	-0.0001 (.005)
Landfill	5E-05 (.004)	4E-05 (.150)	-3E-06 (.890)	3E-05 (.096)	-4E-05 (.072)	3E-06 (.911)	2E-05 (.057)	2E-05 (.252)
Industry A	2E-05 (.097)	7E-05 (.009)	4E-05 (.043)	2E-05 (.168)	5E-05 (.022)	9E-05 (.001)	1E-05 (.094)	2E-05 (.047)
Industry B	-3E-05 (.050)	-4E-05 (.326)	2E-05 (.378)	-3E-05 (.076)	-7E-06 (.818)	5E-06 (.922)	-1E-05 (.274)	-4E-05 (.027)
Park	1E-05 (.368)	-4E-05 (.153)	7E-05 (.003)	-1E-05 (.355)	4E-05 (.118)	0.0001 (.001)	-3E-05 (.002)	-3E-05 (.010)
Railway Station	-9E-06 (.452)	-2E-05 (.474)	-3E-05 (.079)	6E-06 (.598)	-7E-06 (.710)	-1E-05 (.654)	2E-06 (.746)	-2E-05 (.007)
Shops	0.0166 (.104)	-0.0033 (.723)	0.0419 (.000)	0.0088 (.404)	-0.0123 (.231)	-0.0126 (.494)	-0.0184 (.009)	-0.05 (.000)
Primary Schools	0.0311 (.473)	0.0703 (.272)	0.2509 (.000)	0.1983 (.000)	0.1454 (.016)	-0.0183 (.857)	0.0795 (.008)	0.0462 (.360)
Acock's Green	-0.0223 (.514)		-0.1993 (.039)	0.0019 (.971)	0.2876 (.007)	-0.7551 (.001)	0.0313 (.541)	-0.2589 (.000)
Aston	-0.4498 (.000)	-0.128 (.039)	-0.2916 (.000)		0.0831 (.244)	-0.5262 (.185)		
Bartley Green	-0.2054 (.023)		-0.3965 (.000)	0.1717 (.042)	0.2538 (.107)	-0.2635 (.303)	0.0238 (.632)	-0.008 (.888)
Billesley	-0.0051 (.915)		-0.7182 (.000)	0.2415 (.000)	0.3353 (.000)	-0.5066 (.072)	0.1318 (.000)	-0.0924 (.085)
Bournville	-0.0057 (.934)		-0.1458 (.000)	0.3945 (.000)	0.5443 (.000)	-0.2491 (.036)	0.197 (.000)	0.1105 (.051)
Brandwood	0.0132 (.858)		-0.2771 (.003)	0.2458 (.001)	0.2601 (.133)	-0.4844 (.066)	0.1434 (.001)	0.0049 (.934)
Edgbaston	0.3266 (.007)		-0.1068 (.068)	0.4475 (.000)	0.3656 (.000)	base case	0.6194 (.000)	0.2778 (.000)
Erdington	0.0667 (.207)		-0.3228 (.000)	0.0517 (.350)	0.1217 (.137)	-0.6826 (.000)	0.0754 (.065)	-0.1392 (.002)

Variable	Parameter Estimates by Submarket (p-values in brackets)							
	1	2	3	4	5	6	7	8
Fox Hollies	base case		-0.2506 (.003)	0.0718 (.217)	0.1899 (.037)	-0.777 (.000)	0.146 (.025)	-0.1174 (.154)
Hall Green	0.0166 (.709)	0.1437 (.169)	-0.0826 (.259)	0.2278 (.000)	0.2345 (.001)	-0.512 (.001)	0.1403 (.000)	-0.0843 (.006)
Handsworth		-0.1655 (.024)	-0.463 (.022)		0.1022 (.051)	-0.1188 (.416)		
Harborne	0.1227 (.114)		0.0789 (.049)	0.3378 (.000)	0.2699 (.001)	0.1197 (.154)	0.3803 (.000)	0.3613 (.000)
Hodge Hill	-0.0969 (.062)		-0.31 (.001)	0.1301 (.004)	0.0421 (.830)		0.0237 (.615)	-0.2982 (.000)
King's Norton	-0.1735 (.044)		-0.1513 (.080)	0.2682 (.004)	0.0798 (.664)	-0.4269 (.009)	0.1001 (.027)	0.0518 (.432)
Kingsbury	0.0822 (.085)		-0.3038 (.068)	0.1072 (.191)	-0.0894 (.726)	-1.848 (.000)	0.0139 (.813)	-0.1415 (.208)
Kingstanding	-0.1328 (.064)			0.0737 (.322)	-0.2064 (.136)	-0.7292 (.005)	0.0025 (.958)	-0.33 (.001)
Ladywood		0.1721 (.030)	-0.345 (.000)		0.3564 (.000)	-0.1095 (.179)		
Longbridge	-0.1718 (.038)		-0.3594 (.000)	0.3264 (.000)	0.1488 (.551)	-0.0789 (.705)	0.1668 (.000)	0.0891 (.303)
Moseley		-0.0167 (.879)	-0.0678 (.209)	0.332 (.000)	0.351 (.000)	-0.1169 (.165)	0.1867 (.000)	0.1854 (.001)
Nechells	-0.0064 (.937)	0.0488 (.365)	-0.4796 (.000)		0.3636 (.000)			
Northfield	-0.1634 (.092)		-0.2858 (.001)	0.2088 (.018)	0.2526 (.392)	-0.2294 (.136)	0.1353 (.001)	0.0485 (.431)
Oscott	-0.1902 (.030)			-0.0316 (.725)	0.1187 (.538)	-0.6522 (.005)	-0.0051 (.821)	0.0506 (.515)
Perry Barr	-0.1687 (.044)		-0.4621 (.000)	0.029 (.751)	0.0713 (.303)	-0.1732 (.488)	base case	-0.1647 (.011)
Quinton	0.0132 (.877)		0.1214 (.175)	0.2221 (.008)	0.37 (.000)	0.0411 (.799)	0.1596 (.000)	0.098 (.092)
Sandwell		-0.1853 (.021)	-0.5654 (.000)		base case	-0.518 (.000)		
Selly Oak	0.0974 (.196)		base case	0.2335 (.015)	0.4646 (.000)	0.0257 (.750)	0.2089 (.000)	0.0996 (.098)
Shard End	0.0291 (.594)		-0.446 (.098)	base case		-1.072 (.022)	0.0144 (.882)	-0.1895 (.387)
Sheldon	0.0585 (.370)			-0.0241 (.577)	-0.2248 (.279)		-0.0315 (.579)	-0.2461 (.013)
Small Heath	0.0052 (.954)	base case	-0.3182 (.000)	-0.0491 (.670)	0.1685 (.040)			
Soho		-0.3121 (.000)	-0.5092 (.000)		0.1215 (.052)			

Variable	Parameter Estimates by Submarket (p-values in brackets)							
	1	2	3	4	5	6	7	8
Sparkbrook		-0.0282 (.571)	-0.2802 (.015)		0.2359 (.000)			
Sparkhill	-0.8373 (.000)	0.1773 (.000)	-0.2816 (.000)		0.4095 (.000)	0.0602 (.592)		
Stockland Green	-0.0636 (.199)		-0.3551 (.000)	-0.0366 (.586)	0.0205 (.774)	-0.6898 (.000)	0.0384 (.557)	
Sutton Four Oaks			0.1379 (.388)	0.3261 (.001)		-0.7361 (.002)	0.2369 (.000)	0.0931 (.009)
Sutton New Hall	0.0633 (.518)		0.2245 (.072)	0.2504 (.000)	0.484 (.036)	-0.7442 (.001)	0.1583 (.000)	0.0236 (.393)
Sutton Vesey			-0.0166 (.895)	0.1698 (.008)	0.2161 (.330)	-0.8229 (.000)	0.1485 (.000)	base case
Washwood Heath	-0.0917 (.045)	0.0004 (.994)	-0.3987 (.000)	0.1517 (.413)	0.0892 (.193)	-0.5821 (.003)		
Weoley	-0.0862 (.251)		0.0803 (.564)	0.255 (.003)	0.1769 (.348)	0.1859 (.325)	0.1863 (.002)	0.0934 (.203)
Yardley	-0.0486 (.204)		-0.334 (.005)	0.0488 (.285)	-0.0203 (.846)	-0.6791 (.001)	-0.0179 (.722)	-0.2011 (.001)
K	71	48	81	74	83	89	72	79
h	.588	.493	.408	.447	.418	.458	.379	.366
N	1,488	1,017	1,527	1,358	1,058	424	2,333	1,432

Table C2: Parameter Estimates for the Spatial Smoothing Partial Linear Model

Variable	Parameter Estimates by Submarket (p-values in brackets)							
	1	2	3	4	5	6	7	8
Nonparametric Variables:								
ln(Floor Area)	0.287	0.3919	0.3045	0.1957	0.347	0.2932	0.3864	0.4537
ln(Garden Area)	0.0594	0.077	0.0705	0.062	0.1623	0.1719	0.0887	0.1231
Age	-0.009	-0.0059	-0.0082	-0.0054	-0.0208	-0.0046	-0.0076	-0.0218
Sale Date	0.0001	0.0002	0.0001	0.0001	0	0	0.0001	0.0002
Structural Variables:								
Bedrooms 1	0.0473 (.777)	-0.0292 (.906)	0.0453 (.784)	-0.0281 (.692)	-0.0557 (.792)		-0.065 (.726)	-0.4505 (.000)
Bedrooms 2	0.0002 (.993)	0.0073 (.695)	-0.0283 (.025)	-0.0285 (.104)	-0.0356 (.225)	-0.0018 (.982)	-0.0297 (.006)	-0.0448 (.055)
Bedrooms 3	base case	base case	base case	base case	base case	base case	base case	base case
Bedrooms 4	-0.1018 (.196)	0.0405 (.408)	0.0306 (.399)	0.0861 (.001)	0.0628 (.032)	0.0304 (.447)	0.0215 (.534)	0.0509 (.000)
Bedrooms 5		0.0422 (.728)	0.1771 (.056)		0.0958 (.068)	0.013 (.783)	0.2714 (.010)	0.0682 (.014)
Bedrooms 6		-0.2501 (.324)	0.2627 (.236)		0.001 (.991)	0.0076 (.903)		0.172 (.086)
Bedrooms 7					-0.042 (.860)	-0.17 (.069)		0.2138 (.146)
Bedrooms 8						0.0568 (.582)		
Bedrooms 9						-0.262 (.307)		
Bedrooms 10						-0.4596 (.131)		
Bedrooms 11						-0.265 (.355)		
Bedrooms 12						0.4654 (.228)		
WCs 1	base case	base case	base case	base case	base case	base case	base case	base case
WCs 2	0.0336 (.084)	0.0356 (.085)	0.0339 (.017)	0.0234 (.113)	0.0117 (.518)	0.013 (.661)	-0.011 (.262)	-0.0026 (.810)
WCs 3	0.0367 (.828)		0.1113 (.311)	0.0559 (.369)	-0.0365 (.830)	0.1252 (.052)	0.2219 (.033)	0.0317 (.289)
WCs 4				-0.0375 (.841)		-0.1734 (.238)		
WCs 5						0.0119 (.962)		

Variable	Parameter Estimates by Submarket (p-values in brackets)							
	1	2	3	4	5	6	7	8
Floors 2	base case	base case	base case	base case	base case	base case	base case	base case
Floors 3	-0.0317 (.371)	-0.0326 (.781)	-0.0388 (.336)	-0.0747 (.035)	-0.0397 (.431)	-0.1833 (.003)	-0.0377 (.171)	-0.2109 (.000)
Floors 4			-0.1598 (.192)		-0.1974 (.034)	-0.6347 (.000)		-0.4681 (.000)
Floors 5			0.0281 (.923)			-0.9616 (.000)		
Floors 6						-0.0817 (.748)		
Floors 7						0.6845 (.755)		
Garage	0.0434 (.001)	-0.0272 (.479)	0.0636 (.001)	0.068 (.000)	0.071 (.000)	0.0312 (.434)	0.026 (.000)	0.0562 (.000)
Central Heating	0.0837 (.000)	0.0161 (.555)	0.1005 (.002)	0.0797 (.004)	0.0155 (.689)	0.063 (.586)	0.0728 (.002)	0.2123 (.000)
Detached Bungalow	0.4157 (.071)			0.1045 (.020)	0.1052 (.279)	0.112 (.240)	0.089 (.066)	0.1498 (.000)
Semi-Detached Bungalow	-0.1725 (.291)			0.0558 (.125)	0.0071 (.966)		0.1292 (.043)	0.033 (.457)
End Terrace Bungalow	-0.0366 (.840)		-0.3607 (.196)	-0.2221 (.273)			0.219 (.770)	
Terrace Bungalow							0.063 (.678)	0.1464 (.518)
Detached House	-0.0462 (.367)	-0.0832 (.623)	0.051 (.304)	0.1545 (.000)	0.0866 (.003)	0.1606 (.000)	0.124 (.000)	0.1126 (.000)
Semi-Detached House	base case	base case	base case	base case	base case	base case	base case	base case
End Terrace House	-0.0573 (.000)	-0.0841 (.156)	-0.0691 (.008)	-0.1054 (.000)	0.0001 (.997)	-0.2559 (.014)	-0.0766 (.000)	-0.0657 (.023)
Terrace House	-0.0874 (.000)	-0.1046 (.053)	-0.0799 (.001)	-0.0859 (.000)	-0.0056 (.818)	-0.0158 (.806)	-0.0829 (.000)	-0.044 (.104)
BG1			-0.0563 (.765)	0.0929 (.641)			0.0591 (.699)	-0.4009 (.030)
BG2				0.3694 (.052)		-0.0443 (.736)		0.4335 (.018)
BG3	-0.0614 (.269)	-0.0632 (.016)	-0.0375 (.128)	-0.0509 (.423)	0.0019 (.982)		0.0677 (.226)	-0.0262 (.880)
BG4	0.0391 (.289)	base case	base case	0.1163 (.005)	base case	0.089 (.271)	0.018 (.496)	0.1041 (.002)
BG5			-0.129 (.277)		0.0774 (.117)	0.047 (.631)		0.1492 (.054)
BG8	0.0978 (.289)	0.1497 (.077)	0.0148 (.551)	-0.0438 (.621)	0.0644 (.119)	-0.2127 (.118)	0.0746 (.278)	0.0948 (.031)

Variable	Parameter Estimates by Submarket (p-values in brackets)							
	1	2	3	4	5	6	7	8
BG9		-0.1015 (.675)			0.1622 (.038)	-0.0189 (.783)		0.1846 (.000)
BG10						-0.0609 (.439)		0.2077 (.142)
BG19	-0.8608 (.000)		-0.1435 (.117)	0.2484 (.006)	0.1522 (.024)	-0.0598 (.501)	0.1369 (.030)	0.0857 (.003)
BG20	base case	0.0814 (.353)	0.0329 (.419)	-0.0969 (.004)	0.0549 (.170)	-0.2742 (.041)	-0.0688 (.000)	-0.0009 (.979)
BG21	0.0885 (.000)	0.1886 (.026)	0.0363 (.391)	0.0316 (.285)	0.1122 (.005)	-0.0954 (.099)	base case	base case
BG24			-0.1849 (.375)		0.2013 (.026)	base case	0.1914 (.008)	0.1513 (.000)
BG25						0.2308 (.080)		
BG30	0.0364 (.267)	0.0756 (.755)	-0.0452 (.776)	-0.106 (.000)	0.1328 (.041)	-0.269 (.049)	-0.0216 (.298)	-0.0568 (.106)
BG31	0.0701 (.112)	0.0826 (.973)	0.2336 (.004)	base case	0.0771 (.388)	-0.2271 (.039)	0.0406 (.271)	-0.1033 (.039)
BG32	0.0798 (.694)		0.4211 (.000)	0.1032 (.017)	0.2838 (.011)	-0.0036 (.964)	-0.0058 (.904)	-0.019 (.710)
BG35						0.0362 (.805)		0.0271 (.760)
BG36	0.0206 (.858)	0.0901 (.971)	0.0125 (.916)	-0.0825 (.290)	-0.3153 (.027)			-0.2008 (.289)
Neigbourhood Variables:								
Wealth	-0.1032 (.000)	0.001 (.976)	-0.0561 (.002)	-0.1291 (.000)	-0.1137 (.000)	-0.1614 (.000)	-0.0646 (.000)	-0.084 (.000)
Ethnicity	-0.1013 (.001)	0.0479 (.020)	-0.0432 (.018)	-0.1108 (.006)	-0.0452 (.002)	-0.1469 (.000)	-0.0851 (.001)	-0.1001 (.004)
Age	-0.0057 (.722)	0.0248 (.330)	0.0043 (.699)	0.0352 (.002)	0.0452 (.006)	0.0549 (.001)	0.0167 (.045)	0.0611 (.000)
Family	-0.0331 (.004)	-0.0267 (.116)	-0.0313 (.004)	-0.0408 (.001)	-0.0349 (.010)	-0.028 (.181)	-0.01 (.361)	-0.0162 (.258)
Environmental Variables:								
Road Noise	0.0018 (.165)	0.0035 (.070)	-0.0053 (.000)	-0.0028 (.046)	-0.0055 (.002)	-0.0021 (.445)	-0.0018 (.047)	-0.0025 (.038)
Rail Noise	-0.0084 (.081)	-0.0068 (.168)	-0.0063 (.063)	-0.0135 (.001)	-0.005 (.316)	-0.0049 (.729)	0.0001 (.974)	-0.0085 (.039)
Air Noise	-0.016 (.001)		-0.0154 (.279)	0.0032 (.467)	0.0339 (.125)	-0.023 (.904)	-0.0063 (.169)	0.0033 (.819)
Park Views	-1E-04 (.517)	0.0007 (.038)	-0.0001 (.495)	0.0002 (.283)	-6E-05 (.779)	-0.0003 (.231)	2E-05 (.878)	0.0003 (.015)

Variable	Parameter Estimates by Submarket (p-values in brackets)							
	1	2	3	4	5	6	7	8
Water Views	0.0048 (.036)	-0.0188 (.053)	0.0003 (.426)	-0.0086 (.015)	0.0047 (.139)	0.0019 (.375)	-0.0027 (.304)	0.0001 (.805)
Locational Variables:								
CBD	6E-05 (.153)	0.0001 (.260)	-8E-05 (.167)	-8E-05 (.054)	-0.0002 (.000)	-3E-05 (.766)	1E-06 (.963)	-7E-05 (.063)
Airport	-7E-05 (.059)	-0.0001 (.110)	7E-05 (.222)	-7E-05 (.095)	0.0001 (.029)	-4E-05 (.711)	-8E-05 (.003)	-6E-05 (.075)
Landfill	5E-05 (.005)	-6E-05 (.023)	-1E-05 (.686)	4E-05 (.012)	-4E-05 (.026)	1E-05 (.717)	4E-05 (.001)	3E-05 (.058)
Industry A	-1E-05 (.418)	5E-05 (.028)	4E-05 (.031)	1E-05 (.270)	2E-05 (.257)	3E-05 (.308)	1E-05 (.079)	2E-05 (.027)
Industry B	-2E-05 (.321)	0.0001 (.011)	2E-05 (.565)	-4E-05 (.030)	7E-06 (.799)	-6E-05 (.231)	-2E-05 (.065)	-7E-05 (.000)
Park	1E-05 (.444)	-1E-05 (.721)	5E-05 (.025)	-3E-05 (.061)	2E-05 (.444)	0.0002 (.000)	-1E-05 (.348)	-1E-05 (.480)
Railway Station	-1E-05 (.228)	6E-06 (.775)	-4E-05 (.021)	-6E-06 (.545)	-1E-05 (.549)	4E-05 (.186)	-1E-05 (.065)	-2E-05 (.006)
Shops	0.0191 (.046)	-0.0145 (.103)	0.025 (.006)	-0.0053 (.612)	-0.0068 (.502)	-0.0033 (.849)	-0.0332 (.000)	-0.0467 (.000)
Primary Schools	0.1134 (.009)	0.1246 (.049)	0.2895 (.000)	0.2072 (.000)	0.1067 (.077)	0.0764 (.482)	0.0741 (.013)	0.0914 (.073)
K	41	37	46	44	47	57	44	52
h	.694	.551	.574	.918	.603	.752	.547	.538
N	1484	1016	1523	1362	1058	424	2341	1432

Table C3: Average Implicit Price Estimates for the Spatial Constants Partial Linear Model

Variable	Average Implicit Prices by Submarket (£s) (standard deviation of mean below)							
	1	2	3	4	5	6	7	8
Structural Variables:								
Bedrooms 1	2,247 (359)	-2,632 (596)	1,802 (517)	3,049 (806)	-12,543 (4,427)		-4,394 (812)	-46,524 (14,489)
Bedrooms 2	678 (108)	115 (26)	-1,167 (335)	-459 (121)	-1,592 (562)	-98 (56)	-1,770 (327)	-4,350 (1,354)
Bedrooms 3	base case	base case	base case	base case	base case	base case	base case	base case
Bedrooms 4	-1,396 (223)	2,035 (461)	1,581 (454)	2,636 (697)	2,344 (827)	3,798 (2,183)	1,020 (188)	4,556 (1,419)
Bedrooms 5		1,757 (398)	10,210 (2,932)		4,970 (1,754)	2,577 (1,481)	12,890 (2,384)	6,348 (1,977)
Bedrooms 6		-2,462 (557)	7,957 (2,285)		-1,614 (570)	4,692 (2,697)		25,021 (7,792)
Bedrooms 7					-3,217 (1,135)	-1,360 (781)		24,009 (7,477)
Bedrooms 8						-238 (137)		
Bedrooms 9						-23,459 (13,484)		
Bedrooms 10						-84,608 (48,633)		
Bedrooms 11						-35,025 (20,133)		
Bedrooms 12						48,186 (27,697)		
WCs 1	base case	base case	base case	base case	base case	base case	base case	base case
WCs 2	1,217 (194)	805 (182)	1,850 (531)	556 (147)	804 (283)	-3,358 (1,930)	-766 (141)	-1,019 (317)
WCs 3	3,191 (510)		5,007 (1,438)	703 (186)	-387 (136)	9,750 (5,604)	11,515 (2,130)	2,147 (668)
WCs 4				-3,222 (852)		4,261 (2,449)		
WCs 5						444.51 (255.51)		
Floors 2	base case	base case	base case	base case	base case	base case	base case	base case
Floors 3	-1,090 (174)	-875 (198)	42 (12)	-4,870 (1,288)	-3,821 (1,349)	-26,898 (15,461)	-1,247 (230)	-26,495 (8,251)
Floors 4			-8,082 (2,321)		-8,526 (3,010)	-87,842 (50,492)		-52,024 (16,202)
Floors 5			7,079 (2,033)			-133,795 (76,906)		

Variable	Average Implicit Prices by Submarket (£s) (standard deviation of mean below)							
	1	2	3	4	5	6	7	8
Floors 6						-49,420 (28,407)		
Garage	2,211 (353)	-967 (219)	2,959 (850)	3,755 (993)	3,545 (1,251)	1,576 (906)	1,719 (318)	5,853 (1,822)
Central Heating	3,767 (602)	203 (45)	6,946 (1,995)	4,585 (1,213)	3,366 (1,188)	4,177 (2,401)	4,133 (764)	20,179 (6,284)
Detached Bungalow	15,405 (2,462)			4,376 (1,157)	2,731 (964)	17,078 (9,816)	5,468 (1,011)	15,533 (4,837)
Semi-Detached Bungalow	-11,973 (1,913)			1,886 (499)	-2,657 (938)		7,379 (1,365)	3,237 (1,008)
End Terrace Bungalow	-2,072 (331)		-10,234 (2,939)	-12,399 (3,280)			6,610 (1,222)	
Terrace Bungalow							27 (5)	18,125 (5,645)
Detached House	252 (40)	-1,132 (256)	5,600 (1,608)	8,434 (2,231)	3,862 (1,363)	22,684 (13,039)	6,760 (1,250)	10,286 (3,203)
Semi-Detached House	base case	base case	base case	base case	base case	base case	base case	base case
End Terrace House	-2,186 (349)	-3,518 (797)	-1,282 (368)	-5,172 (1,368)	564 (199)	-33,622 (19,326)	-4,314 (798)	-7,154 (2,228)
Terrace House	-3,451 (551)	-4,224 (957)	-2,361 (678)	-3,970 (1,050)	-267 (94)	4,349 (2,500)	-4,399 (813)	-5,669 (1,765)
BG1			931 (267)	2,568 (679)			-5,217 (965)	-28,786 (8,965)
BG2				10,720 (2,836)		-3,896 (2,239)		45,106 (14,047)
BG3	-484 (77)	-1,204 (272)	-880 (252)	-383 (101)	8,129 (2,869)		4,198 (776)	3,635 (1,132)
BG4	592 (94)	base case	base case	5,217 (1,380)	base case	13,226 (7,602)	782 (144)	8,478 (2,640)
BG5			-10,681 (3,068)		4,722 (1,667)	5,878 (3,378)		25,949 (8,081)
BG8	2,515 (402)	4,296 (973)	1,197 (343)	-3,292 (871)	3,815 (1,347)	-41,970 (24,124)	3,620 (669)	10,095 (3,144)
BG9		-5,532 (1,253)			8,047 (2,841)	-3,925 (2,256)		14,071 (4,382)
BG10						-14,362 (8,255)		21,786 (6,784)
BG19	-34,821 (5,565)		2,162 (621)	9,860 (2,608)	4,316 (1,523)	7,007 (4,028)	7,638 (1,413)	8,168 (2,543)
BG20	base case	5,928 (1,343)	854 (245)	-6,587 (1,742)	6,444 (2,275)	-31,116 (17,885)	-3,812 (705)	-275 (85)
BG21	3,548 (567)	6,237 (1,413)	2,344 (673)	617 (163)	7,780 (2,746)	-15,476 (8,896)		

Variable	Average Implicit Prices by Submarket (£s) (standard deviation of mean below)							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
BG24			-15,365 (4,413)		11,337 (4,002)	base case	5,056 (935)	14,286 (4,449)
BG25						22,608 (12,995)	base case	base case
BG30	938 (149)	4,027 (912)	-2,405 (691)	-7,954 (2,104)	8,259 (2,915)	-43,744 (25,144)	-1,876 (347)	-5,753 (1,791)
BG31	2,776 (443)	29,619 (6,710)	17,710 (5,087)	base case	13,210 (4,663)	-4,318 (2,482)	-705 (130)	-16,190 (5,042)
BG32	1,987 (317)		28,478 (8,180)	3,497 (925)	27,121 (9,574)	4,174 (2,399)	-2,005 (371)	-8,468 (2,637)
BG35						25,073 (14,412)		-473 (147)
BG36	-1,376 (220)	25,436 (5,762)	9,880 (2,838)	-5,926 (1,567)	-13,067 (4,613)			-16,383 (5,102)
Neighbourhood Variables:								
Wealth	-4,863 (777)	-141 (31)	-2,179 (625)	-5,811 (1,537)	-5,358 (1,891)	-23,118 (13,288)	-3,028 (560)	-8,780 (2,734)
Ethnicity	-3,803 (607)	1,205 (273)	-2,006 (576)	-5,039 (1,333)	-1,628 (574)	-28,399 (16,324)	-4,431 (819)	-10,392 (3,236)
Age	69 (11)	1,188 (269)	64 (18)	774 (204)	1,951 (688)	9,167 (5,269)	899 (166)	5,806 (1,808)
Family	-1,271 (203)	-974 (220)	-829 (238)	-2,598 (687)	-1,367 (482)	-4,270 (2,454)	-621 (115)	-2,499 (778)
Environmental Variables:								
Road Noise	80 (12)	57 (12)	-229 (66)	-162 (42)	-318 (112)	-484 (278)	-173 (32)	-311 (97)
Rail Noise	-391 (62)	-108 (24)	-422 (121)	-691 (182)	-730 (257)	-1,748 (1,005)	-26 (4)	-766 (238)
Air Noise	-558 (89)		410 (117)	176 (46)	2,125 (750)	4,237 (2,435)	-346 (64)	327 (101)
Park Views	-4.04 (0.64)	15.25 (3.46)	-15.09 (4.33)	10.82 (2.86)	0.48 (0.17)	-14.23 (8.18)	-1.17 (0.22)	14.67 (4.57)
Water Views	252 (40)	-426 (96)	20 (6)	-575 (152)	241 (85)	231 (133)	-138 (25)	0.38 (0.12)
Locational Variables:								
CBD	-0.03 (0.00)	7.82 (1.77)	-6.73 (1.93)	-1.31 (0.35)	-7.87 (2.78)	27.35 (15.72)	0.74 (0.14)	-5.95 (1.85)
Airport	-0.14 (0.02)	-8.65 (1.96)	4.73 (1.36)	-5.15 (1.36)	-0.06 (0.02)	-39.25 (22.56)	-4.29 (0.79)	-10.47 (3.26)
Landfill	1.96 (0.31)	1.35 (0.31)	-0.14 (0.04)	1.46 (0.39)	-1.96 (0.69)	0.44 (0.25)	1.14 (0.21)	1.61 (0.50)
Industry A	0.97 (0.16)	2.21 (0.50)	1.77 (0.51)	0.87 (0.23)	2.42 (0.85)	12.29 (7.06)	0.67 (0.12)	1.60 (0.50)

Variable	Average Implicit Prices by Submarket (£s) (standard deviation of mean below)							
	1	2	3	4	5	6	7	8
Industry B	-1.36	-1.36	1.12	-1.77	-0.35	0.64	-0.76	-3.52
	(0.22)	(0.31)	(0.32)	(0.47)	(0.12)	(0.37)	(0.14)	(1.10)
Park	0.49	-1.36	3.04	-0.69	1.97	17.41	-1.71	-3.17
	(0.08)	(0.31)	(0.87)	(0.18)	(0.70)	(10.01)	(0.32)	(0.99)
Railway Station	-0.35	-0.65	-1.40	0.34	-0.38	-1.69	0.13	-2.30
	(0.06)	(0.15)	(0.40)	(0.09)	(0.13)	(0.97)	(0.02)	(0.72)
Shops	668	-102	1,940	491	-646	-1,722	-1,018	-4,900
	(106)	(23)	(557)	(130)	(228)	(989)	(188)	(1,526)
Primary Schools	1,250	2,176	11,622	11,058	7,620	-2,496	4,413	4,528
	(199)	(493)	(3,338)	(2,925)	(2,690)	(1,434)	(816)	(1,410)
Acock's Green	-897		-9,235	106	15,071	-102,913	1,738	-25,378
	(143)		(2,652)	(28)	(5,320)	(59,155)	(321)	(7,903)
Aston	-18,107	-3,961	-13,512		4,352	-71,714		
	(2,893)	(897)	(3,881)		(1,536)	(41,222)		
Bartley Green	-8,269		-18,371	9,574	13,301	-35,907	1,321	-781
	(1,321)		(5,277)	(2,533)	(4,695)	(20,639)	(244)	(243)
Billesley	-203		-33,276	13,468	17,575	-69,046	7,314	-9,061
	(32)		(9,558)	(3,562)	(6,204)	(39,688)	(1,353)	(2,821)
Bournville	-230		-6,754	21,995	28,524	-33,954	10,932	10,835
	(36)		(1,940)	(5,818)	(10,069)	(19,517)	(2,022)	(3,374)
Brandwood	532		-12,838	13,707	13,630	-66,023	7,960	475
	(85)		(3,687)	(3,626)	(4,811)	(37,951)	(1,472)	(148)
Edgbaston	13,148		-4,947	24,951	19,163	base case	34,375	27,230
	(2,101)		(1,421)	(6,600)	(6,765)		(6,359)	(8,480)
Erdington	2,683		-14,955	2,880	6,377	-93,037	4,185	-13,640
	(428)		(4,295)	(762)	(2,251)	(53,479)	(774)	(4,248)
Fox Hollies	base case		-11,611	4,005	9,950	-105,891	8,102	-11,511
			(3,335)	(1,059)	(3,512)	(60,867)	(1,498)	(3,585)
Hall Green	666	4,446	-3,828	12,701	12,289	-69,780	7,784	-8,259
	(106)	(1,007)	(1,099)	(3,360)	(4,338)	(40,110)	(1,440)	(2,572)
Handsworth		-5,120	-21,451		5,354	-16,188		
		(1,160)	(6,161)		(1,890)	(9,305)		
Harborne	4,939		3,657	18,837	14,143	16,309	21,107	35,416
	(789)		(1,050)	(4,983)	(4,992)	(9,374)	(3,904)	(11,029)
Hodge Hill	-3,900		-14,364	7,252	2,205		1,312	-29,230
	(623)		(4,125)	(1,918)	(778)		(242)	(9,103)
King's Norton	-6,985		-7,010	14,955	4,180	-58,188	5,554	5,078
	(1,116)		(2,013)	(3,956)	(1,475)	(33,447)	(1,027)	(1,581)
Kingsbury	3,307		-14,077	5,979	-4,686	-251,835	769	-13,867
	(528)		(4,043)	(1,581)	(1,654)	(144,757)	(142)	(4,318)
Kingstanding	-5,345			4,110	-10,817	-99,378	136	-32,348
	(854)			(1,087)	(3,818)	(57,123)	(25)	(10,074)
Ladywood		5,326	-15,982		18,677	-14,925		
		(1,206)	(4,590)		(6,593)	(8,579)		
Longbridge	-6,915		-16,653	18,202	7,797	-10,746	9,255	8,730
	(1,105)		(4,783)	(4,815)	(2,752)	(6,177)	(1,712)	(2,718)

Variable	Average Implicit Prices by Submarket (£s) (standard deviation of mean below)							
	1	2	3	4	5	6	7	8
Moseley		-515 (116)	-3,141 (902)	18,510 (4,897)	18,397 (6,494)	-15,928 (9,155)	10,361 (1,916)	18,174 (5,660)
Nechells	-255 (40)	1,509 (341)	-22,221 (6,383)		19,057 (6,727)			
Northfield	-6,578 (1,051)		-13,243 (3,804)	11,643 (3,080)	13,236 (4,672)	-31,266 (17,972)	7,508 (1,388)	4,753 (1,480)
Oscott	-7,658 (1,223)			-1,761 (466)	6,221 (2,196)	-88,885 (51,092)	-285 (52)	4,962 (1,545)
Perry Barr	-6,789 (1,085)		-21,412 (6,150)	1,614 (427)	3,734 (1,318)	-23,607 (13,570)	base case	-16,143 (5,027)
Quinton	530 (84)		5,623 (1,615)	12,386 (3,276)	19,393 (6,846)	5,603 (3,220)	8,855 (1,638)	9,601 (2,990)
Sandwell		-5,733 (1,298)	-26,198 (7,525)		base case	-70,597 (40,579)		
Selly Oak	3,918 (626)		base case	13,022 (3,445)	24,349 (8,595)	3,505 (2,014)	11,595 (2,145)	9,763 (3,040)
Shard End	1,170 (187)		-20,663 (5,935)	base case		-146,086 (83,971)	796 (147)	-18,575 (5,785)
Sheldon	2,355 (376)			-1,343 (355)	-11,781 (4,159)		-1,746 (323)	-24,118 (7,511)
Small Heath	209 (33)	base case	-14,741 (4,234)	-2,735 (723)	8,830 (3,117)			
Soho		-9,658 (2,188)	-23,594 (6,777)		6,367 (2,247)			
Sparkbrook		-872 (197)	-12,983 (3,729)		12,363 (4,364)			
Sparkhill	-33,706 (5,387)	5,485 (1,242)	-13,047 (3,747)		21,460 (7,575)	8,201 (4,714)		
Stockland Green	-2,560 (409)		-16,453 (4,726)	-2,039 (539)	1,073 (378)	-94,007 (54,036)	2,129 (394)	
Sutton Four Oaks			6,388 (1,835)	18,182 (4,810)		-100,318 (57,664)	13,150 (2,432)	9,121 (2,840)
Sutton New Hall	2,547 (407)		10,400 (2,987)	13,961 (3,693)	25,368 (8,955)	-101,429 (58,302)	8,784 (1,625)	2,312 (720)
Sutton Vesey			-768 (220)	9,466 (2,504)	11,324 (3,997)	-112,151 (64,465)	8,240 (1,524)	base case
Washwood Heath	-3,691 (590)	10.90 (2.47)	-18,472 (5,306)	8,460 (2,238)	4,676 (1,650)	-79,328 (45,599)		
Weoley	-3,470 (554)		3,718 (1,068)	14,216 (3,760)	9,273 (3,273)	25,335 (14,563)	10,338 (1,912)	9,157 (2,851)
Yardley	-1,956 (312)		-15,475 (4,445)	2,721 (719)	-1,061 (374)	-92,552 (53,200)	-992 (183)	-19,714 (6,139)

Table C4: Average Implicit Price Estimates for the Spatial Smoothing Partial Linear Model

Variable	Average Implicit Prices by Submarket (£s) (standard deviation of mean below)							
	1	2	3	4	5	6	7	8
Structural Variables:								
Bedrooms 1	1,903 (293)	-901 (207)	2,098 (569)	-1,561 (399)	-2,899 (999)		-3,589 (638)	-44,223 (13,042)
Bedrooms 2	6.21 (0.96)	227 (52)	-1,310 (355)	-1,584 (404)	-1,849 (637)	-240 (131)	-1,638 (291)	-4,394 (1,296)
Bedrooms 3	base case	base case	base case	base case	base case	base case	base case	base case
Bedrooms 4	-4,099 (633)	1,250 (287)	1,415 (384)	4,791 (1,224)	3,267 (1,126)	4,105 (2,246)	1,188 (211)	4,997 (1,473)
Bedrooms 5		1,304 (299)	8,208 (2,227)		4,982 (1,717)	1,757 (961)	14,985 (2,663)	6,689 (1,972)
Bedrooms 6		-7,727 (1,777)	12,174 (3,304)		50 (17)	1,021 (558)		16,879 (4,977)
Bedrooms 7					-2,186 (753)	-22,960 (12,563)		20,984 (6,188)
Bedrooms 8						7,670 (4,196)		
Bedrooms 9						-35,385 (19,361)		
Bedrooms 10						-62,082 (33,969)		
Bedrooms 11						-35,793 (19,584)		
Bedrooms 12						62,860 (34,394)		
WCs 1	base case	base case	base case	base case	base case	base case	base case	base case
WCs 2	1,353 (209)	1,098 (252)	1,569 (426)	1,303 (333)	609 (210)	1,760 (963)	-605 (107)	-256 (75)
WCs 3	1,477 (228)		5,160 (1,400)	3,106 (794)	-1,897 (653)	16,905 (9,250)	12,253 (2,178)	3,114 (918)
WCs 4				-2,084 (532)		-23,423 (12,816)		
WCs 5						1,612 (882)		
Floors 2	base case	base case	base case	base case	base case	base case	base case	base case
Floors 3	-1,277 (197)	-1,008 (231)	-1,796 (487)	-4,152 (1,061)	-2,066 (712)	-24,751 (13,542)	-2,083 (370)	-20,701 (6,105)
Floors 4			-7,405 (2,009)		-10,269 (3,539)	-85,727 (46,906)		-45,945 (13,550)

Variable	Average Implicit Prices by Submarket (£s) (standard deviation of mean below)							
	1	2	3	4	5	6	7	8
Floors 5			1,303 (353)			-129,875 (71,062)		
Floors 6						-11,031 (6,035)		
Floors 7						92,455 (50,588)		
Garage	1,745 (269)	-840 (193)	2,948 (800)	3,780 (966)	3,694 (1,273)	4,214 (2,306)	1,433 (254)	5,512 (1,625)
Central Heating	3,370 (520)	496 (114)	4,657 (1,264)	4,435 (1,133)	803 (276)	8,512 (4,657)	4,021 (714)	20,835 (6,144)
Detached Bungalow	16,737 (2,585)			5,811 (1,485)	5,470 (1,885)	15,121 (8,274)	4,915 (873)	14,699 (4,335)
Semi-Detached Bungalow	-6,946 (1,073)			3,103 (793)	370 (127)		7,136 (1,268)	3,237 (954)
End Terrace Bungalow	-1,474 (227)		-16,714 (4,536)	-12,356 (3,158)			12,095 (2,150)	
Terrace Bungalow							3,480 (618)	14,373 (4,238)
Detached House	-1,858 (287)	-2,569 (591)	2,364 (641)	8,592 (2,196)	4,506 (1,553)	21,696 (11,871)	6,846 (1,217)	11,056 (3,260)
Semi-Detached House	base case	base case	base case	base case	base case	base case	base case	base case
End Terrace House	-2,307 (356)	-2,599 (597)	-3,203 (869)	-5,865 (1,499)	6.28 (2.17)	-34,567 (18,914)	-4,232 (752)	-6,447 (1,901)
Terrace House	-3,520 (543)	-3,231 (743)	-3,701 (1,004)	-4,778 (1,221)	-290 (100)	-2,137 (1,169)	-4,579 (814)	-4,320 (1,274)
BG1			-2,607 (707)	5,169 (1,321)			3,261 (579)	-39,349 (11,604)
BG2				20,546 (5,251)		-5,979 (3,271)		42,552 (12,549)
BG3	-2,470 (381)	-1,953 (449)	-1,738 (471)	-2,832 (724)	98 (34)		3,739 (664)	-2,573 (758)
BG4	1,572 (242)	base case	base case	6,470 (1,653)	base case	12,020 (6,577)	995 (176)	10,222 (3,014)
BG5			-5,977 (1,622)		4,024 (1,387)	6,350 (3,474)		14,647 (4,319)
BG8	3,936 (608)	4,625 (1,063)	685 (186)	-2,435 (622)	3,348 (1,153)	-28,729 (15,719)	4,118 (732)	9,303 (2,743)
BG9		-3,136 (721)			8,438 (2,908)	-2,549 (1,395)		18,116 (5,342)
BG10						-8,224 (4,499)		20,389 (6,013)

Variable	Average Implicit Prices by Submarket (£s) (standard deviation of mean below)							
	1	2	3	4	5	6	7	8
BG19	-34,662		-6,648	13,820	7,917	-8,077	7,562	8,410
	(5,354)		(1,804)	(3,532)	(2,728)	(4,419)	(1,344)	(2,480)
BG20	base case	2,516	1,525	-5,389	2,857	-37,032	-3,800	-88
		(578)	(414)	(1,377)	(984)	(20,262)	(675)	(26)
BG21	3,561	5,828	1,681	1,760	5,836	-12,880	base case	base case
	(550)	(1,340)	(456)	(449)	(2,011)	(7,047)		
BG24			-8,567		10,470	base case	10,571	14,846
			(2,325)		(3,608)		(1,879)	(4,378)
BG25						31,168		
						(17,053)		
BG30	1,463	2,336	-2,094	-5,895	6,906	-36,337	-1,193	-5,573
	(226)	(537)	(568)	(1,506)	(2,380)	(19,882)	(212)	(1,643)
BG31	2,823	2,552	10,827	base case	4,011	-30,677	2,241	-10,140
	(436)	(587)	(2,938)		(1,382)	(16,785)	(398)	(2,990)
BG32	3,211		19,517	5,741	14,761	-482	-319	-1,866
	(496)		(5,297)	(1,467)	(5,087)	(264)	(56)	(550)
BG35						4,883		2,656
						(2,672)		(783)
BG36	831	2,783	577	-4,590	-16,402			-19,708
	(128)	(640)	(156)	(1,173)	(5,653)			(5,812)
Neighbourhood Variables:								
Wealth	-4,154	29	-2,599	-7,180	-5,915	-21,804	-3,568	-8,242
	(641)	(6)	(705)	(1,835)	(2,038)	(11,930)	(634)	(2,430)
Ethnicity	-4,081	1,479	-2,002	-6,162	-2,352	-19,842	-4,697	-9,821
	(630)	(340)	(543)	(1,574)	(810)	(10,856)	(835)	(2,896)
Age	-229	764	197	1,956	2,350	7,410	922	5,999
	(35)	(175)	(53)	(500)	(810)	(4,054)	(163)	(1,769)
Family	-1,332	-823	-1,448	-2,271	-1,813	-3,786	-550	-1,592
	(205)	(189)	(393)	(580)	(624)	(2,072)	(97)	(469)
Environmental Variables:								
Road Noise	74	107	-245	-156	-286	-277	-100	-243
	(11)	(24)	(66)	(39)	(98)	(151)	(17)	(71)
Rail Noise	-338	-210	-291	-749	-260	-662	5.89	-831
	(52)	(48)	(79)	(191)	(89)	(362)	(1.05)	(245)
Air Noise	-643		-714	176	1,762	-3,109	-350	319
	(99)		(193)	(45)	(607)	(1,701)	(62)	(94)
Park Views	-3.90	22.45	-6.63	11.82	-2.99	-36.30	0.91	25.91
	(0.60)	(5.16)	(1.80)	(3.02)	(1.03)	(19.86)	(0.16)	(7.64)
Water Views	191	-579	15	-481	244	255	-150	10.05
	(29)	(133)	(4)	(122)	(84)	(139)	(26)	(2.96)
Locational Variables:								
CBD	2.38	3.25	-3.80	-4.55	-12.40	-4.72	0.07	-6.82
	(0.37)	(0.75)	(1.03)	(1.16)	(4.27)	(2.58)	(0.01)	(2.01)

Variable	Average Implicit Prices by Submarket (£s) (standard deviation of mean below)							
	1	2	3	4	5	6	7	8
Airport	-2.70 (0.42)	-3.85 (0.88)	3.06 (0.83)	-3.79 (0.97)	6.21 (2.14)	-5.60 (3.06)	-4.47 (0.79)	-6.34 (1.87)
Landfill	1.83 (0.28)	-1.79 (0.41)	-0.45 (0.12)	2.21 (0.56)	-2.32 (0.80)	1.66 (0.91)	2.06 (0.37)	2.99 (0.88)
Industry A	-0.39 (0.06)	1.53 (0.35)	1.93 (0.52)	0.77 (0.20)	1.08 (0.37)	4.54 (2.48)	0.72 (0.13)	2.03 (0.60)
Industry B	-0.66 (0.10)	3.21 (0.74)	0.74 (0.20)	-2.11 (0.54)	0.38 (0.13)	-8.56 (4.68)	-1.17 (0.21)	-6.39 (1.89)
Park	0.45 (0.07)	-0.34 (0.08)	2.30 (0.62)	-1.51 (0.39)	0.94 (0.32)	22.49 (12.31)	-0.52 (0.09)	-0.94 (0.28)
Railway Station	-0.55 (0.08)	0.20 (0.05)	-1.73 (0.47)	-0.36 (0.09)	-0.57 (0.19)	5.15 (2.82)	-0.63 (0.11)	-2.39 (0.70)
Shops	769 (118)	-447 (102)	1,157 (314)	-293 (74)	-355 (122)	-447 (245)	-1,830 (325)	-4,580 (1,350)
Primary Schools	4,565 (705)	3,848 (885)	13,416 (3,641)	11,526 (2,946)	5,549 (1,912)	10,313 (5,642)	4,092 (727)	8,974 (2,646)

APPENDIX D: NEIGHBOURHOOD & PROPERTY PARTITION PARAMETER ESTIMATES

Table D1: Parameters of Hedonic Price Equations for Neighbourhood Socioeconomic Characteristics Partition

Variable	**Cluster 1**	**Cluster 2**	**Cluster 3**	**Cluster4**	**Cluster 5**	**Cluster 6**	**Cluster 7**
Constant	8.9067***	8.2327***	8.7047***	8.4688***	8.3764***	8.4239***	8.9403***
Structural Characteristics:							
Floor Area (log)	0.3827***	0.3991***	0.4383***	0.3612***	0.4879***	0.3864***	0.3670***
Garden Area (log)	0.0838***	0.1662***	0.0973***	0.1005***	0.0940***	0.1393***	0.1446***
Garage	0.0448***	0.0579***	0.0550***	0.0350*	0.0524***	0.0607***	0.0369
Central Heating	0.0464*	0.0653*	0.0577**	-0.0283	0.1032***	0.0828**	-0.0716*
Age	-0.0148**	-0.0067	-0.0096	-0.0058	-0.0204***	-0.0091	-0.0106
WCs							
One	b	b	b	b	b	b	b
Two	0.0243*	-0.0399**	0.0315**	-0.0059	0.0297**	-0.022	-0.0244
Three	0.0198	0.2056**	-0.0112	-0.022	0.1304***	0.0222	0.0075
Four	.	0.8627***	-0.2295*	.	0.4666**	.	.
Five	.	0.372	.	.	.	.	.
Bedrooms							
One	0.0727	0.0675	0.0414	0.2473**	0.0351	0.3394	0.1957
Two	0.007	-0.0013	0.0127	-0.0299	0.0152	-0.0062	0.0560**
Three	b	b	b	b	b	b	b

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6	Cluster 7
Four	0.0278	0.0029	0.0165	0.0279	0.0407*	0.0677**	0.047
Five	0.0452	-0.0609	0.1349***	0.1474**	0.1459***	0.1758***	0.044
Six	-0.1925*	-0.0346	0.0692	0.4663*	0.3435***	0.1821**	-0.0534
Seven	0.1105	-0.3969***	0.6348***	-0.2241	-0.5027***	0.2726*	0.4141
Eight	.	-0.0483	.	0.4797**	-0.0843	.	0.2517
Nine	.	.	0.0221	.	.	.	.
Storeys							
One	-0.07	-0.4751	-0.037	0.0449	0.1903***	0.2255	0.1281
Two	b	b	b	b	b	b	b
Three	-0.0481	-0.2195***	-0.1069***	-0.0672	-0.1115***	-0.0166	0.0183
Four	-0.2106*	-0.8875***	-0.4576***	-0.1522	-0.1909*	-0.1956	-0.4995**
Five	.	.	.	-0.3947*	-0.3526	.	-0.5758*
Construction Type							
Detached Bungalow	0.2031	0.8569***	0.1771	0.1213	.	-0.1787	.
Semi-Detached Bungalow	0.1699	.	0.0075	.	-0.0383	.	-0.0694
End Terrace Bungalow	-0.0122	.	.	.	.	.	.
Terrace Bungalow	.	.	.	.	.	.	0.0827
Detached House	0.1396***	0.1477***	0.1220***	0.1386***	0.1087***	0.0721***	0.0884**
Semi-Detached House	b	b	b	b	b	b	b
End Terrace House	-0.0887***	-0.0981***	-0.0440**	-0.0493*	-0.0780***	-0.0309	-0.1012***
Terrace House	-0.0795***	-0.0418*	-0.0647***	-0.0407*	-0.0917***	-0.0763***	-0.0833***

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6	Cluster 7
Beacon Group							
1. Unrenovated cottage pre 1919	-0.1371	.	.	0.0228	-0.3349	-0.0215	.
2. Renovated cottage pre 1919	.	0.1031	0.3737*	.	0.6604***	0.2239	0.2426
3. Small "industrial" pre 1919	-0.1158***	0.0321	0.1031**	0.0414	-0.0325	-0.0665	-0.1759***
4. Medium "industrial" pre 1919	-0.0225	0.0257	-0.0485*	-0.0053	-0.0259	0.0092	-0.0545
5. Large terrace pre 1919	0.0134	0.0614	-0.0662	-0.1176	0.0702	0.029	-0.1293
8. Small "villa" pre 1919	0.0092	0.0927	-0.0518	0.0416	0.0324	-0.0365	0.2134***
9. Large "villas" pre 1919	0.036	0.0955	0.0222	0.1769*	-0.0274	0.1872**	0.1927**
10. Large detached pre 1919	0.4524**	-0.1677	0.2266***	0.0403	0.4721***	-0.1772	-0.4598*
19. Houses 1908 to 1930	0.1012**	0.072	-0.0585	0.0704	0.0748	0.1327**	0.1309*
20. Subsidy houses 1920s & 30s	-0.0805***	-0.0919***	-0.0880***	-0.1278***	-0.0292	0.0157	-0.0545
21. Standard houses 1919-45	b	b	b	b	b	b	b
24. Large houses 1919-45	0.2597***	0.1120**	0.2021***	0.1768**	0.1352***	0.2615***	0.1892**
25. Individual houses 1919-45	0.1885	.	0.0152	.	-0.3234	0.1769	.
30. Standard houses 1945-53	-0.0851***	-0.1605***	-0.1436***	-0.0127	-0.0845***	-0.0851*	-0.0498
31. Standard houses post 1953	-0.0491	-0.0169	-0.0204	-0.0261	-0.0543	0.0304	-0.011
32. Large houses post 1953	0.1536***	0.1201*	0.1202***	0.1581**	0.0157	0.1359**	0.1013
35. Individual houses post 1945	0.5353***	0.4387*	-0.214	0.0809	0.0312	0.0838	.
36. "Town Houses" post 1950	-0.2377***	-0.1489	-0.1329	.	-0.2121	-0.2559***	-0.1947
Sale Date							
1st Quarter (Jan. to Mar.)	-0.0508***	-0.0313*	-0.0564***	-0.0400**	-0.0407***	-0.0716***	-0.0675***

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6	Cluster 7
2nd Quarter (Apr. to June)	-0.0231*	-0.017	-0.0224*	-0.0495***	-0.009	0.0246	-0.0262
3rd Quarter (July to Sept.)	b	b	b	b	b	b	b
4th Quarter (Oct. to Dec.)	-0.0039	-0.0094	0.0023	-0.0171	0.015	0.0269	-0.0101
Neighbourhood Characteristics							
Poverty Factor	-0.0871***	-0.0736***	-0.0648***	-0.1284***	-0.0471***	-0.1260***	-0.0483**
Skills Factor	0.0628***	0.0305***	0.0524***	0.0401**	0.0662***	0.0055	0.0404**
Age Factor	0.0209***	0.0303**	0.014	0.0264*	0.0098	0.0587***	0.0404**
Family Factor	-0.0075	-0.0489***	-0.0205*	-0.0255	-0.0106	-0.0248	-0.0506***
Asian Factor	0.0137	-0.0482***	-0.0368***	-0.0697**	-0.0118	0.0314**	0.0438**
Black Factor	-0.0254**	0.0046	-0.0517***	-0.0314	-0.0520***	0.0269**	-0.011
Locational Characteristics							
Proximity to City Centre	0.0001**	-0.0001	-0.0001**	-0.0001	0	0.0001*	-0.0001*
Proximity and Quantity of Shops	0.0230***	-0.0134	-0.0347***	-0.0023	0.0126	0.0271**	-0.0363***
Proximity and Quality of Primary Schools	0.0961**	0.1593***	0.1089***	0.1766***	0.0942**	0.0134	0.0404
Walking time to Rail Station	0	0	0	0.0001***	0.0000**	0	0
Walking time to a Park	0	0	0.0000**	0	0	0	0
Driving time to Airport	-0.0001***	-0.0001	0	0	-0.0001***	-0.0001**	-0.0001
Proximity to A-Type Industrial Processes	0	0.0001***	0	0.0000**	0.0000**	0	0
Proximity to B-Type Industrial Processes	0	-0.0001***	0	-0.0001*	0	-0.0001***	0

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6	Cluster 7
Proximity to Land Fill sites	0	0.0000**	0.0000**	0.0001***	0.0000*	0.0000**	0
Wards							
Acock's Green	-0.2896**	-0.0878	-0.2595***	0.0285	-0.1522***	-0.128	0.0541
Aston	.	-0.5673	-0.3934**	.	.	-0.2820**	-0.1593
Bartley Green	-0.1273	.	-0.2058***	-0.1212	-0.1503*	.	.
Billesley	-0.0747	-0.0923	-0.1043**	0.0655	-0.0829*	.	.
Bournville	0.0683	0.2489*	-0.0475	0.078	0.0654	.	.
Brandwood	-0.1095	0.1651	-0.1362**	0.0854	0.0599	0.056	.
Edgbaston	0.0399	0.2047	-0.2538***	0.118	0.1255*	0.5427***	0.3242
Erdington	-0.2274**	0.0281	-0.1146***	0.0413	-0.1196***	-0.0675	.
Fox Hollies	-0.2115*	.	-0.2427***	0.1233*	-0.1544***	-0.1841*	0.1992
Hall Green	-0.1784	0.0166	-0.1187***	0.0472	-0.0839**	0.122	0.1605
Handsworth	.	-0.2885**	-0.2472*	.	-0.2132***	-0.2763**	-0.0937
Harborne	0.0748	0.3169**	0.0817	0.1066	0.3960***	.	.
Hodge Hill	-0.2692**	.	-0.2717***	0.1024	-0.0918**	-0.0024	.
King's Norton	-0.0621	.	-0.0571	-0.0436	-0.0245	.	.
Kingsbury	-0.2660**	.	-0.2635***	0.0237	-0.0245	.	.
Kingstanding	-0.1627	-0.1531	-0.1734***	-0.0853	-0.0694	0.0205	-0.2233
Ladywood	-0.084	0.1058	-0.3156***	0.1636	-0.0301	-0.221	.
Longbridge	-0.1124	0.169	-0.1032	-0.0639	0.0824	.	0.4550***
Moseley	.	0.155	-0.3543***	-0.3153**	-0.0606	0.1903	0.3409**

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6	Cluster 7
Nechells	-0.3682***	-0.2002	-0.1422**	0.0916	-0.4287***	-0.7439***	-0.0714
Northfield	-0.0855	.	-0.0881	0.0913	0.1691	.	.
Oscott	-0.2088*	-0.3334**	-0.2240***	-0.2329**	-0.1163**	.	.
Perry Barr	-0.1629	-0.2357*	-0.2433***	-0.1577	-0.1565***	.	.
Quinton	0.0412	0.2236	-0.1776*	-0.0967	0.0335	0.1068	.
Sandwell	-0.1557	-0.3896***	-0.2749***	-0.0042	-0.2002***	-0.0467	-0.1068
Selly Oak	0.0982	0.2363	.	0.2023*	0.1423**	.	.
Shard End	-0.4113***	.	-0.2220***	0.118	-0.1345	-0.4445***	.
Sheldon	-0.2656**	.	-0.2177***	.	0.0045	-0.067	-0.0842
Small Heath	-0.2866**	0.0406	.	.	-0.2047	-0.1810*	-0.1147
Soho	.	-0.3117**	.	.	.	-0.4080***	-0.2027
Sparkbrook	.	0.0503	.	.	.	-0.1595	-0.0112
Sparkhill	-0.1936	0.0833	-0.058	0.2220*	-0.0819	-0.0868	0.2269
Stockland Green	.	.	-0.2485***	-0.0557	-0.1976***	.	.
Sutton Four Oaks	0.0501	-0.1293	0.0659*	0.1483	0.0454	.	.
Sutton New Hall	b	B	b	b	b	b	b
Sutton Vesey	-0.0638	.	.	0.0936	-0.0472	0.1214	0.2678*
Washwood Heath	-0.3167***	.	-0.3352***	0.0473	-0.1505***	-0.2755**	-0.4276
Weoley	-0.11	0.0734	-0.2655***	0.0939	-0.0187	0.1514	0.1722
Yardley	-0.2692**	-0.0777	-0.2097***	-0.0155	.	-0.1086	0.0141
Environmental Characteristics							

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6	Cluster 7
Views of Water	0.0055**	-0.0001	0	0.0029	-0.0009	-0.0008	0.0002
Views of Parkland	0	-0.0002	-0.0002	0	0.0002	0.0003	0
Road Traffic Noise	-0.0004	-0.0002	-0.0024**	-0.0037**	-0.0038***	-0.0035**	-0.0035*
Rail Traffic Noise	-0.0026	-0.0126*	-0.0086**	-0.0089**	-0.0023	-0.0046	-0.0119**
Aircraft Noise	-0.0906*	-0.1413	0.0102	-0.0637	.	.	-0.0109
K	96	90	96	93	97	85	82
N	2261	1258	2173	895	2018	1207	970
R^2	0.721	0.830	0.800	0.807	0.790	0.847	0.829
s^2	0.0455	0.0471	0.0456	0.0382	0.0457	0.0514	0.0588

b Base case for a set of dummy variables
* Significant at 10% level of confidence
** Significant at 5% level of confidence
*** Significant at 1% level of confidence

Table D2: Parameters of Hedonic Price Equations for Property Attributes Partition

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6
Constant	8.9319***	8.6194***	8.4596***	8.0269***	8.2942***	8.5659***
Structural Characteristics:						
Floor Area (log)	0.2795***	0.4408***	0.3860***	0.5532***	0.4595***	0.4090***
Garden Area (log)	0.0822***	0.0659***	0.1256***	0.0766***	0.0587***	0.1667***
Garage	0.0762***	0.0063	0.0349	0.0567***	0.0416***	0.0467**
Central Heating	0.0185	0.0204	-0.008	0.0701**	0.0742***	0.2206***
Age	-0.0556***	-0.0131**	0.0149	-0.0116	-0.01	-0.0281***
WCs						
One	b	b	b	b	b	b
Two	-0.0216	0.0148	0.0268	0.0211	0.0037	-0.0027
Three	0.0207	.	0.1565	.	-0.032	0.0373
Four	-0.0882	.	.	.	.	0.3829
Five	.	.	.	.	.	0.2078
Bedrooms						
One	-0.0434	0.1536	0.8127***	-0.0036	0.104	-0.4211***
Two	-0.0117	-0.0006	-0.0607*	0.0131	0.006	0.0404
Three	b	b	b	b	b	b
Four	0.0552**	0.0891**	0.0114	0.0323	-0.0143	0.0224
Five	0.2768***	0.2126	0.0113	.	0.0546	0.0523*

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6
Six	0.0806	-0.1611	0.0624	.	.	0.0511
Seven	.	.	-0.1037	.	.	-0.0395
Eight	.	.	-0.4149	.	.	-0.0712
Nine	.	.	-0.5133*	.	.	.
Storeys						
One	0.1283	.	-0.3993	0.0474	-0.0201	-0.0447
Two	b	b	b	b	b	b
Three	-0.1098**	0.0676	-0.1193***	-0.0598	-0.0684***	-0.1682***
Four	-0.3953***	.	-0.4555***	.	.	-0.0463
Five	.	.	-0.7086***	.	.	-0.2824
Construction Type						
Detached Bungalow	0.0918	.	.	0.1858**	.	0.1756***
Semi-Detached Bungalow	-0.1722	.	.	.	0.1684**	.
End Terrace Bungalow	-0.2379	.	.	.	.	0.5389**
Terrace Bungalow	.	.	.	.	.	.
Detached House	0.1697***	-0.2648***	0.1527**	0.0836***	0.0611***	0.1185***
Semi-Detached House	b	b	b	b	b	b
End Terrace House	-0.0607***	-0.0531*	-0.036	-0.0818***	-0.0474***	-0.1310***
Terrace House	-0.0760***	-0.0740***	0.0019	-0.0764***	-0.0663***	-0.017
Beacon Group						
1. Unrenovated cottage pre 1919	.	0.0157	0.0739	-0.2473	.	.

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6
2. Renovated cottage pre 1919	.	.	0.0873	0.5262**	.	0.3516***
3. Small "industrial" pre 1919	.	-0.0458***	-0.0887	.	0.2271	0.0294
4. Medium "industrial" pre 1919	b	b	b	b	b	b
5. Large terrace pre 1919	.	-0.3192	0.0271	.	.	0.1018
8. Small "villa" pre 1919	.	0.1162***	0.0486	-0.0165	0.1986	-0.0999
9. Large "villas" pre 1919	.	.	0.0499	.	.	0.1009*
10. Large detached pre 1919	.	.	-0.2982*	.	.	-0.0102
19. Houses 1908 to 1930	.	0.0965	0.1787***	0.1524	0.1675**	-0.0186
20. Subsidy houses 1920s & 30s	.	0.0524	-0.0571	0.0032	0.1088**	-0.1699***
21. Standard houses 1919-45	.	0.1853***	-0.0717	0.0716	0.2073***	-0.1203***
24. Large houses 1919-45	.	.	-0.0338	0.4530***	0.2834***	0.0077
25. Individual houses 1919-45	.	.	.	.	.	-0.2203
30. Standard houses 1945-53	-0.287	.	.	-0.0253	0.1	-0.2545***
31. Standard houses post 1953	0.5236***	.	.	-0.0163	.	-0.2048***
32. Large houses post 1953	0.5639***	.	.	-0.2800**	0.4193***	-0.107
35. Individual houses post 1945	0.6099***	.	.	.	.	0.1258
36. "Town Houses" post 1950	0.3640***	.	.	-0.1128	.	.
Sale Date						
1st Quarter (Jan. to Mar.)	-0.0435***	-0.0756***	-0.0626**	-0.0447***	-0.0363***	-0.0555***
2nd Quarter (Apr. to June)	-0.0147	-0.0276**	0.0142	-0.0222	-0.0097	-0.0404**
3rd Quarter (July to Sept.)	b	b	b	b	b	b

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6
4th Quarter (Oct. to Dec.)	-0.0016	-0.0038	0.0283	0.0241	0.0039	-0.0237
Neighbourhood Characteristics						
Poverty Factor	-0.1069***	-0.0496***	-0.0719***	-0.1023***	-0.0756***	-0.0163
Skills Factor	0.0078	0.0469***	0.0742***	0.0261**	0.0293***	0.0582***
Age Factor	0.0173**	0.0261**	0.0385**	0.0126	0.0370***	0.0522***
Family Factor	-0.011	-0.0386***	-0.0074	-0.0035	-0.0051	-0.0084
Asian Factor	-0.011	0.0326***	-0.0441**	0.0316	0.0072	-0.0681**
Black Factor	-0.0500***	0.0190**	-0.0346**	-0.0217	-0.0431***	-0.0632***
Locational Characteristics						
Proximity to City Centre	0	0.0001	-0.0002	0	0	-0.0001**
Proximity and Quantity of Shops	0.0048	0.0115*	0.0117	-0.0168	0.0105*	-0.0350***
Proximity and Quality of Primary Schools	0.1033**	0.0726*	0.2103**	0.1804***	0.0897***	0.0252
Walking time to Rail Station	0.0000**	0.0000**	0	0.0000**	0.0000***	0.0000**
Walking time to a Park	0	0	0.0001*	0.0000*	0	0
Driving time to Airport	-0.0001**	0	-0.0001	-0.0001***	-0.0001***	-0.0001**
Proximity to A-Type Industrial Processes	0	0.0000***	0.0001**	0	0.0000**	0.0000***
Proximity to B-Type Industrial Processes	0	0	0	0	0.0000***	0
Proximity to Land Fill sites	0.0000**	0	-0.0001*	0.0001***	0	0.0000***
Wards						

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6
Acock's Green	-0.0985**	-0.2980***	-0.1735	-0.2238***	-0.1338***	-0.3068***
Aston	-0.1998***	-0.4744***	-0.2	-0.2945*	-0.3373***	.
Bartley Green	-0.0526	-0.5340**	.	-0.1370*	-0.2129***	-0.1053
Billesley	-0.1131**	-0.6512***	0.2093	-0.0018	-0.0607	-0.0484
Bournville	0.0687	-0.0984	0.1722	0.0723	0.0207	0.1982**
Brandwood	-0.0394	-0.1043	-0.0247	0.0477	-0.056	0.0407
Edgbaston	0.1398**	0.0486	0.0294	0.1335	0.3624*	0.0787
Erdington	-0.1848***	-0.3205***	-0.1721	-0.1381***	-0.0773	-0.1442***
Fox Hollies	-0.1697*	-0.3561***	-0.2723**	-0.2320***	-0.0689	-0.0807
Hall Green	-0.0787	-0.2242***	-0.1633	-0.1694***	-0.0329	-0.0237
Handsworth	-0.2831***	-0.5141***	-0.2026	-0.0984	-0.1776***	-0.1319
Harborne	0.104	0.1831**	0.3256**	0.0684	0.0666	0.3190***
Hodge Hill	-0.2725***	-0.2894***	-0.5603*	-0.0917*	-0.1560***	-0.0930*
King's Norton	-0.0365	0.0785	0.195	-0.0082	-0.0606	0.1458
Kingsbury	-0.0874*	0.0114	0.1579	-0.2104***	-0.1471***	-0.0318
Kingstanding	-0.0694	-0.6128**	0.16	-0.3544***	-0.2166***	-0.2141**
Ladywood	-0.1048	-0.2405***	0.0595	0.0529	0.0634	-0.0252
Longbridge	0.0882	.	0.2941	0.0402	-0.084	0.1962*
Moseley	-0.1062	-0.0388	0.0472	0.0526	-0.0069	0.0964
Nechells	-0.0133	-0.4448***	0.1268	-0.0549	-0.0976	-0.3924**
Northfield	-0.0606	-0.1584*	0.1384	0.0003	-0.0464	0.0858

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6
Oscott	-0.1006	.	.	-0.0478	-0.2653***	-0.0749
Perry Barr	-0.1452**	-0.4767***	.	-0.1237*	-0.1902***	-0.2123**
Quinton	-0.0711	0.0591	0.3211*	0.0047	-0.0732	0.089
Sandwell	-0.0959	-0.5222***	-0.3727***	-0.1515**	-0.2528***	-0.2675***
Selly Oak	0.0689	0.0528	0.0966	0.0408	0.0666	0.1404*
Shard End	-0.2729***	.	-0.2846	-0.2451***	-0.2070***	-0.0312
Sheldon	-0.3771***	0.0495	.	-0.2283***	-0.1800***	-0.1985**
Small Heath	-0.6870***	-0.3508***	-0.0682	-0.1042	-0.1147*	-0.0814
Soho	-0.2091***	-0.5691***	-0.3294**	-0.2014	-0.2353***	0.1125
Sparkbrook	-0.0528	-0.3781***	-0.107	.	.	0.2386
Sparkhill	0.0673	-0.2401***	-0.0182	0.0984	0.1890*	0.0937
Stockland Green	-0.1466**	-0.4050***	-0.2686**	-0.1826***	-0.1295***	-0.2040***
Sutton Four Oaks	0.1189**	-0.0879	0.0258	0.0998*	-0.0204	0.0198
Sutton New Hall	b	b	b	b	b	b
Sutton Vesey	-0.054	-0.0491	-0.0267	0.0435	-0.0474	-0.0377
Washwood Heath	.	-0.4153***	-0.1691	-0.3485***	-0.1846***	0.0246
Weoley	-0.0086	-0.3608***	0.1447	0.0239	-0.1183**	0.1101
Yardley	-0.1448**	-0.3017***	-0.1195	-0.2531***	-0.1809***	-0.1586***
Environmental Characteristics						
Views of Water	-0.0023	0	0.0003	-0.0029	0.001	0.0004
Views of Parkland	-0.0002	0	0.0001	-0.0001	0	0

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6
Road Traffic Noise	-0.0024	-0.0022**	-0.0052***	-0.0019	-0.0016*	-0.0017
Rail Traffic Noise	-0.0063*	-0.0074**	-0.0055	-0.0128**	-0.0042	-0.0039
Aircraft Noise	0.0092	.	.	-0.0072	-0.0123	-0.0095
K	88	80	91	87	86	101
N	1540	2324	878	1176	3453	1353
R^2	0.760	0.646	0.655	0.686	0.574	0.763
s^2	0.0387	0.0525	0.0772	0.038	0.0355	0.0486

b Base case for a set of dummy variables
* Significant at 10% level of confidence
** Significant at 5% level of confidence
*** Significant at 1% level of confidence

APPENDIX E: NO LOCATIONAL CONSTANTS MODEL PARAMETER ESTIMATES

Table E1: No Locational Constants Model parameter estimates

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6	Cluster 7
Constant	8.4851***	7.8679***	8.1762***	8.3128***	7.6377***	7.9238***	8.2710***
Structural Characteristics:							
Floor Area (log)	0.3804***	0.4254***	0.4623***	0.4199***	0.5455***	0.4195***	0.4221***
Garden Area (log)	0.0828***	0.1681***	0.1043***	0.0911***	0.0894***	0.1498***	0.1318***
Garage	0.0503***	0.0579***	0.0527***	0.0430**	0.0457***	0.0451**	0.0313
Central Heating	0.0239	0.0013	0.0752***	-0.033	0.1386***	0.0909**	-0.0819**
Age	-0.005	0.0001	-0.0127*	-0.008	0.0035	0.0021	-0.0098
WCs							
One	b	b	b	b	b	b	b
Two	0.0221	-0.0397**	0.0454***	-0.0147	0.0334**	-0.0317	-0.0214
Three	0.0386	0.1677	-0.0138	0.0452	0.1056**	-0.0661	0.0139
Four	.	0.6197*	-0.2030*	.	0.4440*	.	.
Five	.	0.1736	.	.	.	.	.
Bedrooms							
One	0.0062	-0.0855	0.033	0.2299*	0.0658	0.1125	0.2077
Two	0.0067	-0.005	0.0056	-0.0297	0.0154	-0.0196	0.0599**
Three	b	b	b	b	b	b	b

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6	Cluster 7
Four	0.0319	0.0039	0.0008	0.0219	0.0336	0.0797***	0.0556
Five	0.0077	-0.0303	0.0905*	0.1743**	0.1161**	0.2038***	0.0619
Six	-0.2325**	-0.0362	-0.0106	0.3477	0.3703***	0.1266	0.034
Seven	0.0924	-0.4692***	0.5032**	-0.0941	-0.4015**	0.163	0.4468*
Eight	.	-0.1253	.	0.4468*	-0.0352	.	0.3906
Nine	.	.	-0.1555	.	.	.	.
Storeys							
One	-0.0273	-0.2675	-0.0095	0.1076*	0.1931***	0.046	0.0904
Two	b	b	b	b	b	b	b
Three	-0.0582	-0.2184***	-0.0880**	-0.055	-0.0975***	-0.0336	0.0179
Four	-0.1976*	-0.9372***	-0.4144***	-0.1468	-0.1842	-0.306	-0.6345***
Five	.	.	.	-0.4005*	-0.29	.	-0.7839**
Construction Type							
Detached Bungalow	0.1623	0.6566*	0.1512***	0.0712	.	.	.
Semi-Detached Bungalow	0.1663	.	.	.	0.0407	-0.0427	-0.1108
End Terrace Bungalow	.	.	.	.	.	.	.
Terrace Bungalow	0.0305	.	0.0082	.	.	.	0.0862
Detached House	0.1594***	0.1717***	0.1227***	0.1479***	0.1089***	0.0536*	0.0682
Semi-Detached House	b	b	b	b	b	b	b
End Terrace House	-0.0891***	-0.1017***	-0.0331*	-0.0436*	-0.0668***	-0.0024	-0.1009***
Terrace House	-0.0720***	-0.0355	-0.0634***	-0.0367	-0.0840***	-0.0642**	-0.1032***

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6	Cluster 7
Beacon Group							
1. Unrenovated cottage pre 1919	-0.2055	.	.	0.0791	-0.3116	0.3689	.
2. Renovated cottage pre 1919	.	0.0174	0.3770*	.	0.6351***	0.1352	0.3305
3. Small “industrial” pre 1919	-0.1129***	0.0886	0.1503***	0.1013*	-0.0799*	-0.1381**	-0.1991***
4. Medium “industrial” pre 1919	-0.0153	0.0186	0.0098	0.0209	-0.0707**	-0.0562	-0.0763
5. Large terrace pre 1919	-0.0142	0.0804	0.0043	-0.0029	-0.0375	-0.0203	-0.0861
8. Small “villa” pre 1919	-0.0482	0.1066*	-0.0196	0.0595	-0.0452	-0.1389	0.1326*
9. Large “villas” pre 1919	0.0346	0.1315*	0.039	0.1847*	-0.0535	0.2349***	0.1346
10. Large detached pre 1919	0.3787*	0.018	0.2219***	0.0779	0.2661*	-0.0541	-0.3021
19. Houses 1908 to 1930	0.0965**	0.0218	-0.0121	0.093	0.014	0.1491**	0.0644
20. Subsidy houses 1920s & 30s	-0.0596***	-0.0533	-0.0647***	-0.0899***	-0.0161	-0.0101	-0.0847*
21. Standard houses 1919-45	b	b	b	b	b	b	b
24. Large houses 1919-45	0.2196***	0.0732	0.2029***	0.1632**	0.1537***	0.2395***	0.088
25. Individual houses 1919-45	0.2847	.	0.1857	.	-0.2497	0.0835	.
30. Standard houses 1945-53	-0.0626**	-0.1043**	-0.1269***	-0.0027	-0.0527*	-0.1485***	-0.0686
31. Standard houses post 1953	0.0101	0.0857*	-0.0049	-0.0013	0.0461	0.0622	0.0098
32. Large houses post 1953	0.2506***	0.2504***	0.1400***	0.1629**	0.1100**	0.1340**	0.1578*
35. Individual houses post 1945	0.5985***	0.411	-0.1745	0.0926	0.0893	0.0523	.
36. “Town Houses” post 1950	-0.1964***	-0.2505	-0.1535	.	-0.0365	-0.2151**	-0.2501*
Sale Date							
1st Quarter (Jan. to Mar.)	-0.0552***	-0.0351*	-0.0588***	-0.0357*	-0.0358**	-0.0686***	-0.0684***

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6	Cluster 7
2nd Quarter (Apr. to June)	-0.0186	-0.0064	-0.019	-0.0545***	-0.0078	0.0396**	-0.0246
3rd Quarter (July to Sept.)	b	b	b	b	b	b	b
4th Quarter (Oct. to Dec.)	-0.0084	0.0049	0.0034	-0.0284	0.0203	0.0399**	-0.0221
Neighbourhood Characteristics							
Poverty Factor	-0.0816***	-0.0445***	-0.0637***	-0.1070***	-0.0186	-0.1373***	-0.0825***
Skills Factor	0.1009***	0.1026***	0.0901***	0.0741***	0.1376***	0.0652***	0.1060***
Age Factor	0.0181**	0.0336***	0.0188**	0.0246**	0.0154*	0.0788***	0.024
Family Factor	-0.0458***	-0.0967***	-0.0349***	-0.0457***	-0.0212**	-0.0565***	-0.0984***
Asian Factor	-0.0216**	0.001	-0.0167	-0.022	-0.0652***	0.0423***	0.0779***
Black Factor	-0.0321***	-0.0245*	-0.0793***	-0.0497**	-0.0797***	0.0039	0.0139
Locational Characteristics							
Proximity to City Centre	0	0	0	0	0	0.0001*	0
Proximity and Quantity of Shops	0.0280***	-0.0252***	-0.0359***	-0.0014	0.011	0.0166	-0.0133
Proximity and Quality of Primary Schools	0.1402***	0.1734***	0.1188***	0.1412***	0.1274***	0.0826	0.1290**
Walking time to Rail Station	0	0	0.0000**	0	0	0	0
Walking time to a Park	0	0.0000**	0.0000**	0	0	0.0000*	0
Driving time to Airport	0	0.0000***	0	0.0000*	0	0.0000*	0.0000*
Proximity to A-Type Industrial Processes	0	0	0.0000*	0.0000***	0	0.0000**	0.0000**
Proximity to B-Type Industrial Processes	0.0000**	0.0000**	0	0	-0.0001***	0	0

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6	Cluster 7
Proximity to Land Fill sites	0.0000**	0.0000*	0.0000***	0.0001***	0.0000***	0	0.0000*
Environmental Characteristics							
Views of Water	0.0058**	-0.0012	-0.0002	-0.0005	-0.0011	-0.0034	0.0005
Views of Parkland	0	-0.0003*	-0.0002	0.0002	0	0.0004	0
Road Traffic Noise	-0.0004	0.0014	-0.0019*	-0.0044***	-0.0038***	-0.0028*	-0.0013
Rail Traffic Noise	-0.0029	-0.0071	-0.0057	-0.0140***	-0.0045	-0.0052	-0.0142**
Aircraft Noise	-0.0596	-0.1672	-0.0018	-0.0662	.	.	-0.0121
K	64	63	63	61	63	60	62
N	2261	1258	2173	895	2018	1207	970
R^2	0.694	0.788	0.783	0.789	0.759	0.819	0.801
$Adj.R^2$	0.685	0.777	0.776	0.771	0.751	0.810	0.787
s	0.222	0.2394	0.2207	0.2015	0.227	0.2441	0.2587

b Base case for a set of dummy variables
* Significant at 10% level of confidence
** Significant at 5% level of confidence
*** Significant at 1% level of confidence

APPENDIX F: SMOOTH SPATIAL EFFECTS MODEL PARAMETER ESTIMATES

Table F1: SSE parameter estimates

Variable	**Cluster 1**	**Cluster 2**	**Cluster 3**	**Cluster4**	**Cluster 5**	**Cluster 6**	**Cluster 7**
Structural Characteristics:							
Floor Area (log)	0.3627***	0.3321***	0.4327***	0.3594***	0.4662***	0.3757***	0.3711***
Garden Area (log)	0.0829***	0.1672***	0.0935***	0.1051***	0.0836***	0.1353***	0.1329***
Garage	0.0402***	0.0752***	0.0450***	0.0152	0.0366***	0.0602***	0.0516**
Central Heating	0.0378	0.0491	0.0540*	-0.0329	0.1259***	0.0734*	-0.0919**
Age	-0.0136*	-0.0065	-0.0094	0.01	-0.0186**	-0.0098	-0.0164
WCs							
One	b	b	b	b	b	B	b
Two	0.0187	-0.0309**	0.0374***	-0.0177	0.0258**	-0.0206	-0.0322
Three	0.0583	0.1399	-0.0199	-0.0446	0.0768*	-0.0203	-0.0577
Four	.	0.9636***	-0.2579**	.	0.192	.	.
Five	.	0.394	.	.	.	.	.
Bedrooms							
One	0.0574	-0.0211	0.0541	0.2716**	0.0733	0.2203	-0.025
Two	0.0094	-0.0128	0.0217	-0.0105	0.0092	-0.0017	0.0533**
Three	b	b	b	b	b	b	b
Four	0.0389*	0.0339	0.0203	0.0278	0.0553**	0.0552**	0.0458

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6	Cluster 7
Five	0.0780*	-0.0236	0.0970**	0.1448*	0.1599***	0.1564***	0.0479
Six	-0.2504**	-0.0125	0.0698	0.1617	0.2793***	0.0808	-0.1179
Seven	0.1054	-0.2475*	0.3495	-0.1789	-0.4030***	0.1444	0.3967
Eight	.	0.0858	.	0.3422	0.0013	.	0.1812
Nine	.	.	-0.1097	.	.	.	.
Storeys							
One	-0.111	0.3490***	-0.0582	0.0566	0.0949	0.0578	0.3675
Two	b	b	b	b	b	b	b
Three	-0.0353	-0.2155***	-0.1097***	-0.0656	-0.0776**	-0.0206	0.0047
Four	-0.165	-0.8355***	-0.3608***	-0.0745	-0.1747	-0.1914	-0.5552***
Five	.	.	.	-0.4045**	-0.234	.	-0.6024*
Construction Type							
Detached Bungalow	0.1694	.	0.2022***	0.0677	0.099	.	-0.2827
Semi-Detached Bungalow	0.1563	-0.5017	.	.	.	0.126	-0.2649
End Terrace Bungalow	-0.0094	.	.	.	.	.	.
Terrace Bungalow	.	.	0.1702	.	.	.	.
Detached House	0.1367***	0.1173***	0.1223***	0.1149***	0.1130***	0.0698***	0.1110***
Semi-Detached House	b	b	b	b	b	b	b
End Terrace House	-0.1051***	-0.0736***	-0.0495***	-0.0524**	-0.0777***	-0.0136	-0.1303***
Terrace House	-0.0944***	-0.0379*	-0.0567***	-0.0585**	-0.1012***	-0.0589**	-0.1154***
Beacon Group							

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6	Cluster 7
1. Unrenovated cottage pre 1919	-0.082	.	.	0.0164	-0.3308	-0.2744	.
2. Renovated cottage pre 1919	.	0.19	0.4129**	.	0.6484***	0.164	0.15
3. Small "industrial" pre 1919	-0.1401***	0.021	0.0944**	-0.0055	-0.0289	-0.1207*	-0.1614**
4. Medium "industrial" pre 1919	-0.0422	0.0117	-0.0502*	-0.0531	-0.0572*	-0.0184	-0.0265
5. Large terrace pre 1919	-0.0189	0.0599	-0.0677	-0.1966	0.0968	0.053	-0.1649
8. Small "villa" pre 1919	0.0048	0.0389	-0.0408	-0.0003	-0.0311	-0.098	0.2514***
9. Large "villas" pre 1919	0.0611	0.0915	0.045	0.1042	0.0016	0.2684***	0.2272**
10. Large detached pre 1919	0.3673*	-0.1821	0.1573*	0.025	0.0887	0.2041	-0.3399
19. Houses 1908 to 1930	0.1376***	0.0639	-0.0587	0.0744	0.0956*	0.1181*	0.2039**
20. Subsidy houses 1920s & 30s	-0.0714***	-0.0495	-0.0538**	-0.1217***	-0.0232	0.0242	-0.0165
21. Standard houses 1919-45	b	b	b	b	b	b	b
24. Large houses 1919-45	0.2578***	0.1332***	0.2098***	0.2744***	0.1103**	0.2848***	0.1516*
25. Individual houses 1919-45	0.2856	.	0.3533	.	-0.0949	0.2182	.
30. Standard houses 1945-53	-0.1016***	-0.1186**	-0.1212***	-0.0009	-0.0923***	-0.0569	-0.0764
31. Standard houses post 1953	-0.0262	0.0189	-0.0067	0.0188	-0.0661*	0.0036	-0.0748
32. Large houses post 1953	0.1095**	0.1046*	0.1254***	0.1948**	-0.0057	0.0992	0.0498
35. Individual houses post 1945	0.4179***	0.273	-0.4998**	0.1312	0.0148	0.1124	.
36. "Town Houses" post 1950	-0.1819**	-0.0425	-0.0281	.	-0.1648	-0.2992***	-0.2767**
Sale Date							
1st Quarter (Jan. to Mar.)	-0.0534***	-0.0382**	-0.0569***	-0.0414**	-0.0482***	-0.0703***	-0.0705***
2nd Quarter (Apr. to June)	-0.0280**	-0.0226	-0.0209*	-0.0422**	-0.0205*	0.025	-0.022

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6	Cluster 7
3rd Quarter (July to Sept.)	b	b	b	b	b	b	b
4th Quarter (Oct. to Dec.)	-0.0006	-0.0083	-0.005	-0.0189	0.0167	0.0279	-0.0162
Neighbourhood Characteristics							
Poverty Factor	-0.0881***	-0.0430**	-0.0646***	-0.1480***	-0.0396**	-0.1096***	-0.0323
Skills Factor	0.0210*	0.0118	0.0230*	0.0301	0.0305**	0.0131	0.0105
Age Factor	0.0133	0.0353**	0.0086	-0.0199	0.0185	0.0537**	0.012
Family Factor	0.0083	-0.0560***	-0.0206	-0.0412	0.0076	-0.0365*	-0.0505**
Asian Factor	-0.0031	-0.0377*	-0.0028	0.0027	-0.0516	0.0272	0.0236
Black Factor	-0.0032	-0.0132	0.0006	0.0179	-0.0450*	0.0157	-0.0025
Locational Characteristics							
Proximity to City Centre	0.0002**	0	0	-0.0002	0	0.0001	0
Proximity and Quantity of Shops	0.0184	-0.0006	-0.0287**	-0.0237	0.0101	0.0139	-0.0073
Proximity and Quality of Primary Schools	0.1566***	0.0999	0.1460***	0.1481*	0.1498***	0.018	0.1217
Walking time to Rail Station	0	0	0	0	0	-0.0001	0
Walking time to a Park	0	0	0	0	0.0001*	0	0
Driving time to Airport	-0.0002**	-0.0001	-0.0001	0.0002	0	0	-0.0001
Proximity to A-Type Industrial Processes	0	0.0001*	0	0.0001	0	0	0.0002
Proximity to B-Type Industrial Processes	0	0.0001	0	0	0	-0.0001	-0.0001
Proximity to Land Fill sites	0	0.0001	0.0001*	0.0001	0.0001	0	0

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6	Cluster 7
Wards							
Acock's Green	1.5582	-1.2035	0.617	-2.6396	-3.2907	2.0884	-30.4681
Aston	.	0.0857	-0.0436	.	.	2.7161	4,647.448
Bartley Green	1.078	.	2.7871	-3.6536	-3.9907	.	.
Billesley	1.432	-1.0255	1.14	-2.67	-3.4591	.	.
Bournville	1.4856	-0.9024	1.1647	-3.0443	-3.3406	.	.
Brandwood	1.3945	-0.9102	1.0962	-2.8444	-3.2232	1.1456	.
Edgbaston	0.2833	-1.3627	1.7834	-1.9179	-3.6108	0.389	-26.2427
Erdington	1.634	0.6445	0.2054	-0.2892	-0.6302	2.4806	.
Fox Hollies	1.4164	.	1.922	-2.5354	-3.4242	1.9649	-29.8137
Hall Green	1.419	0.02	2.0962	-2.676	-3.335	2.1434	.
Handsworth	.	-0.6367	-0.5662	.	-0.75	2.7422	4,647.558
Harborne	0.9861	-1.0371	2.4572	-3.8086	-3.7257	.	.
Hodge Hill	1.6503	.	0.1127	-2.6101	-1.0371	4.3444	.
King's Norton	1.3624	.	1.1851	-3.1194	-2.7419	.	.
Kingsbury	1.8315	.	0.074	0.2226	-0.578	.	.
Kingstanding	1.3201	0.7698	0.3603	-0.5309	-0.214	3.5735	4,647.727
Ladywood	.	-1.2317	1.7089	-5.5353	-3.5308	2.59	.
Longbridge	1.4468	-0.6193	1.3655	-3.5282	-1.7353	.	-15323.77
Moseley	.	-1.2053	1.4861	-2.836	-3.6227	1.822	-26.4718
Nechells	1.5333	-0.9836	0.0469	-2.9615	-0.8741	2.1878	4,645.471

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6	Cluster 7
Northfield	1.4186	.	1.227	-3.6237	-2.2074	.	.
Oscott	1.3238	0.8529	0.6805	-1.0281	-0.0447	.	.
Perry Barr	1.5363	1.1032	0.6431	-1.138	-0.2238	.	.
Quinton	1.2933	-0.9474	2.187	-3.7711	-3.6272	-27.3819	.
Sandwell	2.0835	-0.1035	0.9237	0.0848	-0.2761	3.1684	4,646.502
Selly Oak	1.4658	-0.9677	.	-3.1897	-3.2682	.	.
Shard End	1.5772	.	0.1567	-2.0472	-17.6935	4.931	.
Sheldon	1.4995	.	0.6312	.	-4.4502	2.0716	-30.7874
Small Heath	1.7	-1.0106	.	.	-0.7566	2.0433	-27.7914
Soho	.	-0.1601	.	.	.	2.6591	4,647.521
Sparkbrook	.	-1.0953	.	.	.	1.7668	-27.3963
Sparkhill	1.7329	-1.0899	2.138	-2.2247	-3.1955	2.149	-26.8326
Stockland Green	.	.	0.3969	0.1092	-0.5616	.	.
Sutton Four Oaks	.	.	0.1087*	0.0299	-0.1252	.	.
Sutton New Hall	b	b	b	b	b	b	b
Sutton Vesey	1.8219	.	.	-0.2182	-0.7973	2.061	4,648.265
Washwood Heath	1.605	.	-0.0751	-2.7345	-0.891	3.4125	-26.2216
Weoley	1.1882	-1.1806	2.255	-3.462	-3.2382	-33.5121	.
Yardley	1.5651	-1.2513	0.6089	-2.6128	.	2.0667	-30.1466
Environmental Characteristics							
Views of Water	0.0044	0.0004	0.0003	0.0021	0.0007	0.0004	0.0002

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster4	Cluster 5	Cluster 6	Cluster 7
Views of Parkland	0	-0.0001	-0.0002	0.0001	-0.0001	0.0004	0
Road Traffic Noise	0.0002	0.0015	-0.0033***	-0.0028	-0.0041***	-0.0022	-0.003
Rail Traffic Noise	-0.0039	-0.0125*	-0.0069*	-0.0087**	0.0011	-0.0052	-0.0103*
Aircraft Noise	-0.0094	-0.024	0.0136	-0.1072	.	.	-0.0143
K	93	88	95	92	96	84	79
N	2261	1258	2173	895	2018	1207	970
R^2	0.760	0.864	0.827	0.838	0.834	0.868	0.853
s^2	0.039	0.038	0.039	0.032	0.036	0.044	0.050
b	550	675	600	650	450	625	400
h	200	300	300	100	200	200	100
Moran's I	-0.016	-0.021	-0.015	-0.048	-0.016	-0.025	-0.025
Probability of Moran's I	0.987	0.953	0.932	0.961	0.915	0.956	0.941

b Base case for a set of dummy variables
* Significant at 10% level of confidence
** Significant at 5% level of confidence
*** Significant at 1% level of confidence

APPENDIX G: WELFARE ESTIMATES

Table G1: Welfare estimates for changes in Road noise exposure at different percentiles of the ethnicity factor distribution

Noise Change (dB)	Welfare Values for Changes in Road Noise at Percentiles of Ethnicity Factor Distribution (£ per year)						
	5th	10th	25th	50th	75th	90th	95th
56 to 55	28.23	28.64	29.36	30.32	32.00	37.08	40.23
57 to 56	30.61	31.02	31.74	32.71	34.39	39.46	42.62
58 to 57	33.00	33.41	34.13	35.09	36.77	41.85	45.01
59 to 58	35.38	35.80	36.52	37.48	39.16	44.24	47.39
60 to 59	37.77	38.18	38.90	39.86	41.54	46.62	49.78
61 to 60	40.16	40.57	41.29	42.25	43.93	49.01	52.16
62 to 61	42.54	42.95	43.67	44.64	46.32	51.39	54.55
63 to 62	44.93	45.34	46.06	47.02	48.70	53.78	56.94
64 to 63	47.31	47.73	48.45	49.41	51.09	56.17	59.32
65 to 64	49.70	50.11	50.83	51.79	53.47	58.55	61.71
66 to 65	52.09	52.50	53.22	54.18	55.86	60.94	64.09
67 to 66	54.47	54.88	55.60	56.57	58.25	63.32	66.48
68 to 67	56.86	57.27	57.99	58.95	60.63	65.71	68.87
69 to 68	59.25	59.66	60.38	61.34	63.02	68.10	71.25
70 to 69	61.63	62.04	62.76	63.72	65.40	70.48	73.64
71 to 70	64.02	64.43	65.15	66.11	67.79	72.87	76.02
72 to 71	66.40	66.81	67.54	68.50	70.18	75.25	78.41
73 to 72	68.79	69.20	69.92	70.88	72.56	77.64	80.80
74 to 73	71.18	71.59	72.31	73.27	74.95	80.03	83.18
75 to 74	73.56	73.97	74.69	75.65	77.33	82.41	85.57
76 to 75	75.95	76.36	77.08	78.04	79.72	84.80	87.95
77 to 76	78.33	78.74	79.47	80.43	82.11	87.18	90.34
78 to 77	80.72	81.13	81.85	82.81	84.49	89.57	92.73
79 to 78	83.11	83.52	84.24	85.20	86.88	91.96	95.11
80 to 79	85.49	85.90	86.62	87.58	89.26	94.34	97.50

Table G2: Welfare estimates for changes in Rail noise exposure at different percentiles of the ethnicity factor distribution

Noise Change (dB)	Welfare Values for Changes in Rail Noise at Percentiles of Ethnicity Factor Distribution) (£ per year)						
	5th	10th	25th	50th	75th	90th	95th
56 to 55	82.05	82.25	82.59	83.05	83.85	86.27	87.77
57 to 56	84.29	84.49	84.83	85.29	86.09	88.51	90.01
58 to 57	86.54	86.73	87.07	87.53	88.33	90.75	92.25
59 to 58	88.78	88.97	89.32	89.77	90.57	92.99	94.50
60 to 59	91.02	91.21	91.56	92.02	92.82	95.23	96.74
61 to 60	93.26	93.46	93.80	94.26	95.06	97.48	98.98
62 to 61	95.50	95.70	96.04	96.50	97.30	99.72	101.22
63 to 62	97.74	97.94	98.28	98.74	99.54	101.96	103.46
64 to 63	99.99	100.18	100.52	100.98	101.78	104.20	105.70
65 to 64	102.23	102.42	102.77	103.22	104.02	106.44	107.95
66 to 65	104.47	104.66	105.01	105.46	106.27	108.68	110.19
67 to 66	106.71	106.91	107.25	107.71	108.51	110.93	112.43
68 to 67	108.95	109.15	109.49	109.95	110.75	113.17	114.67
69 to 68	111.19	111.39	111.73	112.19	112.99	115.41	116.91
70 to 69	113.43	113.63	113.97	114.43	115.23	117.65	119.15
71 to 70	115.68	115.87	116.22	116.67	117.47	119.89	121.40
72 to 71	117.92	118.11	118.46	118.91	119.72	122.13	123.64
73 to 72	120.16	120.36	120.70	121.16	121.96	124.38	125.88
74 to 73	122.40	122.60	122.94	123.40	124.20	126.62	128.12
75 to 74	124.64	124.84	125.18	125.64	126.44	128.86	130.36
76 to 75	126.88	127.08	127.42	127.88	128.68	131.10	132.60
77 to 76	129.13	129.32	129.67	130.12	130.92	133.34	134.84
78 to 77	131.37	131.56	131.91	132.36	133.16	135.58	137.09
79 to 78	133.61	133.81	134.15	134.61	135.41	137.82	139.33
80 to 79	135.85	136.05	136.39	136.85	137.65	140.07	141.57

Table G3: Welfare estimates for changes in Road noise exposure at different percentiles of the Age Factor distribution

Noise Change (dB)	Welfare Values for Changes in Road Noise at Percentiles of Age Factor Distribution (£ per year)						
	5th	10th	25th	50th	75th	90th	95th
56 to 55	37.06	35.77	33.40	31.17	29.30	27.83	26.86
57 to 56	39.45	38.16	35.78	33.56	31.69	30.22	29.25
58 to 57	41.83	40.54	38.17	35.94	34.08	32.60	31.63
59 to 58	44.22	42.93	40.56	38.33	36.46	34.99	34.02
60 to 59	46.61	45.31	42.94	40.72	38.85	37.37	36.40
61 to 60	48.99	47.70	45.33	43.10	41.23	39.76	38.79
62 to 61	51.38	50.09	47.71	45.49	43.62	42.15	41.18
63 to 62	53.76	52.47	50.10	47.87	46.01	44.53	43.56
64 to 63	56.15	54.86	52.49	50.26	48.39	46.92	45.95
65 to 64	58.54	57.24	54.87	52.65	50.78	49.30	48.33
66 to 65	60.92	59.63	57.26	55.03	53.16	51.69	50.72
67 to 66	63.31	62.02	59.64	57.42	55.55	54.08	53.11
68 to 67	65.69	64.40	62.03	59.80	57.94	56.46	55.49
69 to 68	68.08	66.79	64.42	62.19	60.32	58.85	57.88
70 to 69	70.47	69.17	66.80	64.58	62.71	61.23	60.26
71 to 70	72.85	71.56	69.19	66.96	65.09	63.62	62.65
72 to 71	75.24	73.95	71.57	69.35	67.48	66.01	65.04
73 to 72	77.63	76.33	73.96	71.73	69.87	68.39	67.42
74 to 73	80.01	78.72	76.35	74.12	72.25	70.78	69.81
75 to 74	82.40	81.10	78.73	76.51	74.64	73.16	72.19
76 to 75	84.78	83.49	81.12	78.89	77.02	75.55	74.58
77 to 76	87.17	85.88	83.50	81.28	79.41	77.94	76.97
78 to 77	89.56	88.26	85.89	83.66	81.80	80.32	79.35
79 to 78	91.94	90.65	88.28	86.05	84.18	82.71	81.74
80 to 79	94.33	93.03	90.66	88.44	86.57	85.09	84.12

Table G4: Welfare estimates for changes in Rail noise exposure at different percentiles of the Age Factor distribution

Noise Change (dB)	Welfare Values for Changes in Rail Noise at Percentiles of Age Factor Distribution (£ per year)						
	5th	10th	25th	50th	75th	90th	95th
56 to 55	73.25	75.66	80.06	84.20	87.67	90.41	92.21
57 to 56	75.50	77.90	82.31	86.44	89.92	92.65	94.45
58 to 57	77.74	80.14	84.55	88.69	92.16	94.89	96.70
59 to 58	79.98	82.38	86.79	90.93	94.40	97.14	98.94
60 to 59	82.22	84.62	89.03	93.17	96.64	99.38	101.18
61 to 60	84.46	86.87	91.27	95.41	98.88	101.62	103.42
62 to 61	86.70	89.11	93.51	97.65	101.12	103.86	105.66
63 to 62	88.95	91.35	95.75	99.89	103.36	106.10	107.90
64 to 63	91.19	93.59	98.00	102.14	105.61	108.34	110.15
65 to 64	93.43	95.83	100.24	104.38	107.85	110.59	112.39
66 to 65	95.67	98.07	102.48	106.62	110.09	112.83	114.63
67 to 66	97.91	100.32	104.72	108.86	112.33	115.07	116.87
68 to 67	100.15	102.56	106.96	111.10	114.57	117.31	119.11
69 to 68	102.40	104.80	109.20	113.34	116.81	119.55	121.35
70 to 69	104.64	107.04	111.45	115.58	119.06	121.79	123.60
71 to 70	106.88	109.28	113.69	117.83	121.30	124.03	125.84
72 to 71	109.12	111.52	115.93	120.07	123.54	126.28	128.08
73 to 72	111.36	113.76	118.17	122.31	125.78	128.52	130.32
74 to 73	113.60	116.01	120.41	124.55	128.02	130.76	132.56
75 to 74	115.84	118.25	122.65	126.79	130.26	133.00	134.80
76 to 75	118.09	120.49	124.90	129.03	132.51	135.24	137.05
77 to 76	120.33	122.73	127.14	131.28	134.75	137.48	139.29
78 to 77	122.57	124.97	129.38	133.52	136.99	139.73	141.53
79 to 78	124.81	127.21	131.62	135.76	139.23	141.97	143.77
80 to 79	127.05	129.46	133.86	138.00	141.47	144.21	146.01

Table G5: Welfare estimates for changes in Road noise exposure at different percentiles of the Family Factor distribution

Noise Change (dB)	Welfare Values for Changes in Road Noise at Percentiles of Family Factor Distribution (£ per year)						
	5th	10th	25th	50th	75th	90th	95th
56 to 55	27.65	28.61	30.01	31.56	33.03	34.27	35.18
57 to 56	30.04	31.00	32.39	33.94	35.42	36.66	37.56
58 to 57	32.42	33.38	34.78	36.33	37.80	39.05	39.95
59 to 58	34.81	35.77	37.16	38.72	40.19	41.43	42.34
60 to 59	37.20	38.15	39.55	41.10	42.58	43.82	44.72
61 to 60	39.58	40.54	41.94	43.49	44.96	46.20	47.11
62 to 61	41.97	42.93	44.32	45.87	47.35	48.59	49.49
63 to 62	44.35	45.31	46.71	48.26	49.73	50.98	51.88
64 to 63	46.74	47.70	49.09	50.65	52.12	53.36	54.27
65 to 64	49.13	50.08	51.48	53.03	54.51	55.75	56.65
66 to 65	51.51	52.47	53.87	55.42	56.89	58.13	59.04
67 to 66	53.90	54.86	56.25	57.80	59.28	60.52	61.42
68 to 67	56.28	57.24	58.64	60.19	61.66	62.91	63.81
69 to 68	58.67	59.63	61.02	62.58	64.05	65.29	66.20
70 to 69	61.06	62.01	63.41	64.96	66.44	67.68	68.58
71 to 70	63.44	64.40	65.80	67.35	68.82	70.06	70.97
72 to 71	65.83	66.79	68.18	69.73	71.21	72.45	73.35
73 to 72	68.21	69.17	70.57	72.12	73.59	74.84	75.74
74 to 73	70.60	71.56	72.95	74.51	75.98	77.22	78.13
75 to 74	72.99	73.94	75.34	76.89	78.37	79.61	80.51
76 to 75	75.37	76.33	77.73	79.28	80.75	81.99	82.90
77 to 76	77.76	78.72	80.11	81.66	83.14	84.38	85.28
78 to 77	80.14	81.10	82.50	84.05	85.52	86.77	87.67
79 to 78	82.53	83.49	84.88	86.44	87.91	89.15	90.06
80 to 79	84.92	85.87	87.27	88.82	90.30	91.54	92.44

Table G6: Welfare estimates for changes in Rail noise exposure at different percentiles of the Family Factor distribution

Noise Change (dB)	Welfare Values for Changes in Rail Noise at Percentiles of Family Factor Distribution (£ per year)						
	5th	10th	25th	50th	75th	90th	95th
56 to 55	81.83	82.27	82.92	83.64	84.32	84.90	85.32
57 to 56	84.07	84.51	85.16	85.88	86.56	87.14	87.56
58 to 57	86.31	86.75	87.40	88.12	88.80	89.38	89.80
59 to 58	88.55	89.00	89.64	90.36	91.05	91.62	92.04
60 to 59	90.79	91.24	91.89	92.60	93.29	93.86	94.28
61 to 60	93.04	93.48	94.13	94.85	95.53	96.10	96.52
62 to 61	95.28	95.72	96.37	97.09	97.77	98.35	98.77
63 to 62	97.52	97.96	98.61	99.33	100.01	100.59	101.01
64 to 63	99.76	100.20	100.85	101.57	102.25	102.83	103.25
65 to 64	102.00	102.45	103.09	103.81	104.50	105.07	105.49
66 to 65	104.24	104.69	105.33	106.05	106.74	107.31	107.73
67 to 66	106.49	106.93	107.58	108.30	108.98	109.55	109.97
68 to 67	108.73	109.17	109.82	110.54	111.22	111.80	112.21
69 to 68	110.97	111.41	112.06	112.78	113.46	114.04	114.46
70 to 69	113.21	113.65	114.30	115.02	115.70	116.28	116.70
71 to 70	115.45	115.90	116.54	117.26	117.95	118.52	118.94
72 to 71	117.69	118.14	118.78	119.50	120.19	120.76	121.18
73 to 72	119.94	120.38	121.03	121.75	122.43	123.00	123.42
74 to 73	122.18	122.62	123.27	123.99	124.67	125.25	125.66
75 to 74	124.42	124.86	125.51	126.23	126.91	127.49	127.91
76 to 75	126.66	127.10	127.75	128.47	129.15	129.73	130.15
77 to 76	128.90	129.35	129.99	130.71	131.39	131.97	132.39
78 to 77	131.14	131.59	132.23	132.95	133.64	134.21	134.63
79 to 78	133.39	133.83	134.48	135.19	135.88	136.45	136.87
80 to 79	135.63	136.07	136.72	137.44	138.12	138.69	139.11

REFERENCES

Abelson, P. W. (1979). "Property prices and the value of amenities", *Journal of Environmental Economics and Management*, 6, pp 11-28.

Abraham, J. M., W. N. Goetzmann and S. M. Wachter (1994). "Homogenous groupings of metropolitan housing markets", Journal of Housing Economics, 3 (3), pp. 186-206.

Allen, G. R. (1980). *Relationship between highway noise, noise mitigation and resident property values*, FHWA/VA-81/1. Charlottesville, Va.: Virginia Highways and Transportation Research Council.

Allen, M. T., T. M. Springer and N. G. Waller (1995). "Implicit pricing across residential rental markets", *Journal of Real Estate Finance and Economics*, 11, pp 137-151.

Anderson, L. M., and H. K. Cordell (1988). "Influence of trees on residential property values in Athens, Georgia (U.S.A.): A survey based on actual sales prices." *Landscape and Urban Planning*, 15(1-2), pp 153-164.

Anderson, R. J., and D. E. Wise (1977). *The effects of highway noise and accessibility on residential property values*, DOT-FH-11-8841. Springfield, Va.: National Technical Information Service.

Anglin, P. M., and R. Gencay (1996). "Semiparametric estimation of a hedonic price function", *Journal of Applied Econometrics*, 11, pp 633-648.

Anselin, L,. (1988). "Lagrange multiplier test diagnostics for spatial dependence and spatial heterogeneity", *Geographical Analysis*, 20, pp 1-17.

Anselin, L., (1993). "Discrete space autoregressive models" pp 454-68 in M. F. Goodchild, B. O. Parks and L. T. Steyaert (eds), *Environmental Modelling with GIS*, Oxford University Press, Oxford.

Anselin, L., and S. Hudak, (1992). "Spatial econometrics in practice: A review of software options", *Regional Science and Urban Economics*, 2, pp 509-536.

Anselin, L., and S. Rey, (1991). "Properties of tests for spatial dependence in linear regression models", *Geographical Analysis*, 23, pp 112-131.

Arguea, N. M., and C. Hsiao, (1993). "Econometric issues of estimating hedonic price functions", *Journal of Econometrics*, 56, pp 243-267.

Bailey, M. J. (1977). *Report on pilot study: Highway noise and property values*, unpublished paper, University of Maryland.

Bajari, P., and M. E. Khan (2001). "Why do blacks live in the cities and whites live in the suburbs?", *Stanford and Tufts University Working Paper.*

Bajari, P., and M. E. Khan (2002). "Estimating housing demand with an application to explaining racial segregation in cities", *Stanford and Tufts University Working Paper.*

Ball, M. J. and R. M. Kirwan (1977). "Accessibility and supply constraints in the urban housing market", *Urban Studies*, 14, pp 11-32.

Banfield, J. D., and A. E. Raftery (1993). "Model-based Gaussian and non-Gaussian clustering", *Biometrics*, 49, pp 803-821.

Barde, J-P., and D. W. Pearce, (eds.) (1991). *Valuing the Environment*, Earthscan, London.

Bartels, C. P. A., and L. Hordijk (1977). "On the power of the generalized Moran contiguity coefficient in testing for spatial autocorrelation among regression disturbances", *Regional Science and Urban Economics*, 7, pp 83-101.

Bartik, T. J., (1988). "Measuring the benefits of amenity improvements", *Land Economics*, 64(2), pp 172-183.

Basu, S., and T. G. Thibodeau (1998). "Analysis of spatial autocorrelation in house prices", *Journal of Real Estate Finance and Economics*, 17(1), pp 61-85.

Bateman, I. J., B. H. Day, I. Lake and A. A. Lovett (2001). *The Effect of Road Traffic on Residential Property Value: A Literature Review and Hedonic Pricing Study*, Edinburgh: Scottish Executive and The Stationary Office.

Bell, K., and N. E. Bockstael (2000). "Applying the generalised method of moments approach to spatial problems involving micro-level data", *Review of Economics and Statistics*, 82(1), pp 72-82.

Bensmail, H., G. Celeux, A. E. Raftery, and C. P. Robert (1997). "Inference in model-based cluster analysis", *Statistics and Computing*, 7, pp 1-10.

Bertrand, N. F. (1997). "Meta-analysis of studies of willingness to pay to reduce traffic noise", unpublished MSc dissertation, University College London.

Biernacki, C., and G. Govaert (1999). "Choosing models in model-based clustering and discriminant analysis", *Journal of Statistical Computation and Simulation*, 64(1), pp. 49-71.

Biernacki, C., G. Celeux and G. Govaert (1999). "An improvement of the NEC criterion for assessing the number of clusters in a mixture model", *Pattern Recognition Letters*, 20(3), pp. 267-272.

Biernacki, C., G. Celeux and G. Govaert (2003). "Choosing starting values for the EM algorithm for getting the highest likelihood in multivariate Gaussian mixture models", *Computational Statistics and Data Analysis*, 41 (3-4), pp. 561-575.

Blackley, P., J. R. Follain, Jr., and J. Ondrich (1984). "Box-Cox estimation of hedonic models: How serious is the iterative OLS variance bias?", *Review of Economics and Statistics*, 66, pp. 348-353.

Blamire, P.A., & Barnsley, M.J. (1996). Inferring urban land use from an analysis of the spatial and morphological characteristics of discrete scene objects. *In Proceedings of the 22nd Annual conference and exhibition of the Remote Sensing Society.*

Bourassa, S.C., F. Hamelink, M. Hoesli and B. D. MacGregor (1999). "Defining housing submarkets", *Journal Of Housing Economics*, 8, pp. 160-183.

Bowman, A. W., and A. Azzalini (2003). "Computational aspects of nonparametric smoothing with illustrations from the sm library", *Computational Statistics and Data Analysis*, 42, pp. 545-560.

Box, G. E. P. and D. R. Cox (1964) "An analysis of transformation", *Journal of the Royal Statistical Society*, 26, pp 211-243.

Boyle, K. J., P. J. Poor and L. O. Taylor (1999). "Estimating the demand for protecting freshwater lakes from eutrophication", *American Journal of Agricultural Economics*, 81(5), pp. 1118-22.

Brandsma, A. S., and R. H. Ketellapper (1979). "Further evidence on alternative procedures for testing of spatial autocorrelation among regression residuals", in *Exploratory and Explanatory Statistical Analysis of Spatial Data*, C. P. A. Bartels and R. H. Ketellapper (eds.), Martinius Nijhoff, Boston, MA.

Brown, J. N., and H. S. Rosen (1982). "On the estimation of structural hedonic price models", *Econometrica*, 50(3), pp 765-768.

Brown, W., (1994). "Dying from too much dust", *New Scientist*, 12 March, pp 12-13.

Burridge, P., (1980). "Testing for spatial autocorrelation among regression residuals", *Environment and Planning A*, 13, pp 795-800.

Butler, R. V., (1980). "Cross-sectional variation in the hedonic relationship for urban housing markets", *Journal of Regional Science*, 20, pp 439-453.

Byers, S. D., and A. E. Raftery (1998). "Nearest neighbour clutter removal for estimating features in a point process", *Journal of the American Statistical Association*, 93, pp 577-584.

Calthrop, E. J., (1995). *The external costs of road transport fuel: Should the fiscal stance towards diesel be altered?*, unpublished MSc dissertation, University College London.

Campbell, J. G., C. Fraley, D. Stanford, F. Murtagh and A. E. Raftery (1999). "Model-based methods for textile fault detection", *International Journal of Imaging Systems And Technology*, 10 (4), pp. 339-346.

Can, A., (1990). "The measurement of neighbourhood dynamics in urban house prices", *Economic Geography*, 66(3), pp 254-272.

Can, A., (1992). "Specification and estimation of hedonic housing price models", *Regional Science and Urban Economics*, 22, pp 453-474.

Can, A., and I. Megbolugbe (1997). "Specification and estimation of hedonic housing price models", *Journal of Real Estate Finance and Economics*, 14, pp 203-222.

Cassel, E., and R. Mendelsohn (1985). "The choice of functional forms for hedonic price equations: Comment", *Journal of Urban Economics*, 18, pp 135-142.

Celeux, G., G. and Govaert (1995). "Gaussian parsimonious clustering models", *Pattern Recognitio*n, 28, pp. 781–793.

Cheshire, P. and S. Sheppard (1998). "Estimating the demand for housing, land and neighbourhood characteristics", *Oxford Bulletin of Economics and Statistics*, 60(3), pp357-382.

Cicchetti, C. J. and Freeman, A. M. (1971). "Optimal demand and consumer's surplus: further comment", *Quarterly Journal of Economics*, 85:528-539.

Clapp, J. M. (2004). "A semiparametric method for estimating local house price indices", *Real Estate Economics*, 32, pp 127-160.

Cliff, A., and J. K. Ord (1972). "Testing for spatial autocorrelation among regression residuals", *Geographical Analysis* 4, pp. 267-284.

Cliff, A., and J. K. Ord (1973). *Spatial Autocorrelation*, Pion Limited: London.

Collins, A., and A. Evans (1994). "Aircraft noise and residential property values: An artificial neural network approach", *Journal of Transport Economics and Policy*, May, pp 175-97.

Crawley, R. W., (1973). "A case study of the effects of an airport on land values", *Journal of Transport Economics and Policy*, May, pp 144-52.

Cropper, M. L., L. B. Deck and K. E. McConnell (1988). "On the choice of functional form for hedonic price functions", *Review of Economics and Statistics*, 70, pp 668-75.

Dasgupta, A., and A. E. Raftery (1998). " Detecting features in spatial point processes with clutter via model-based clustering", *Journal of the American Statistical Association*, 93, pp. 294-302.

Davidson, R., and J. G. MacKinnon (1981). "Several tests for model specification in the presence of alternative hypotheses", *Econometrica*, 49, pp. 781-793.

Davidson, R., and J. G. MacKinnon (1993). *Estimation and Inference in Econometrics*, Oxford University Press: New York.

Day, B. H., (2003). "Submarket Identification in Property Markets: A Hedonic Housing Price Model for Glasgow", *CSERGE Working Paper*, EDM 03-09, University of East Anglia, UK.

Day, B. H., and I. J. Bateman and I. Lake (2003). "What Price Peace? A Comprehensive Approach to the Specification and Estimation of Hedonic Housing Price Models", *CSERGE Working Paper*, EDM 03-08, University of East Anglia, UK.

De Vany, A. S., (1976). "An economic model of airport noise pollution in an urban environment", in *Theory and Measurement of Economic Externalities*, S. A. Y. Lin (ed.) pp 205-214, Academic Press, New York.

Deaton, A. S., and J. Muellbauer (1980). "An almost ideal demand system", *American Economic Review*, 70, pp 312-326.

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977). "Maximum likelihood for incomplete data via the EM algorithm", *Journal of the Royal Statistical Society B*, 39, pp. 1–38.

Department of the Environment (1972). *Urban Motorways*, The Urban Motorways Project Team, HMSO, London.

Department of the Environment (1987). *Handling Geographic Information*, Report of the Committee of Enquiry chaired by Lord Chorley, HMSO, London.

Department of the Environment Transport and the Regions (2000). *A report on the production of noise maps of the City of Birmingham*. London: HMSO.

Department of Transport (1991). *Transport statistics Great Britain 1991*. London: HMSO.

Department of Transport (1993). *Vehicle speeds in Great Britain 1992*. London: Transport Statistics Report, Department of Transport.

Department of Transport (Welsh Office) (1988). Calculation of Road Traffic Noise. London: HMSO.

Dubin, R. A., (1988). "Estimation of regression coefficients in the presence of spatially autocorrelated error terms", *Review of Economics and Statistics*, 70, pp 466-474.

Dubin, R. A., (1992). "Spatial autocorrelation and neighbourhood quality", *Regional Science and Urban Economics*, 22, pp 433-452.

Dubin, R. A., (1998). "Predicting house prices using multiple listings data", *Journal of Real Estate Finance and Economics*, 17(1), pp 35-59.

Dunteman, G.H. (1994) "Principal components analysis", pp.157-245 in M.S. Lewis-Beck (ed.) *Factor Analysis & Related Techniques*, International Handbooks of Quantitative Applications in the Social Sciences, 5, Sage Publications, London.

Dygert P. K. (1973). *Estimation of the cost aircraft noise to residential activities*, unpublished PhD dissertation, University of Michigan.

Edmonds, R. G., Jr. (1985). "Some evidence on the inter-temporal stability of hedonic price functions", *Land Economics*, 61, pp 445-51.

Ekeland, I., J. J. Heckman and L. Nesheim (2002). "Identifying hedonic models", *American Economic Review*, 92(2), pp. 304-309.

Ekeland, I., J. J. Heckman and L. Nesheim (2003). "Identification and estimation of hedonic models", *Journal of Political Economy*, forthcoming.

Ellner, S. P., and Y. Seifu (2002). "Using spatial statistics to select model complexity", *Journal of Computational and Graphical Statistics*, 11(2), pp. 348-369.

Emerson F. C. (1972). "Valuation of residential amenities: An econometric approach", *Appraisal Journal*, 40, pp 268-278.

Epanechnikov, V., (1969). "Nonparametric estimates of a multivariate probability density", *Theory of Probability and its Applications,* 14, pp 153-158.

Epple, D., (1987). "Hedonic prices and implicit markets: Estimating demand and supply functions for differentiated products", *Journal of Political Economy,* 95, pp 59-80.

Epple, D., and G. Platt (1998). "Equilibrium and Local Redistribution in an Urban Economy when Households Differ in Preferences and Incomes", *Journal of Urban Economics*, 43(1), pp. 23-51

Epple, D., and H. Sieg (1999a). "Estimating Equilibrium Models of Local Jurisdictions", *Journal of Political Economy*, 99(4). Pp.828-858.

Epple, D., and H. Sieg (1999b). "The Tiebout Hypothesis and Majority Rule: An Empirical Analysis," unpublished working paper, Carnegie Mellon University and NBER.

Epple, D., and H. Sieg (2002). "Identification of Equilibrium Models of Local Jurisdictions", unpublished working paper, Carnegie Mellon University and NBER.

Everitt, B. S., (1993). *Cluster Analysis*, 3rd Ed., Halsted Press, London, New York.

Fan, J., (1992). "Design-adaptive nonparameteric regression", *Journal of the American Statistical Association*, 87, pp. 998-1004.

Fan, J., and J. S. Marron (1994). "Fast implementations of nonparametric curve estimators", *Journal of Computational and Graphical Statistics*, 3(1), pp. 35-56.

Fasulo, D., (1999). "An analysis of recent work on clustering algorithms", Technical Report No. 01-03-02, Department of Computer Science and Engineering, University of Washington, Seattle, WA.

Fovell, R. G., (1997). "Consensus Clustering of U.S. Temperature and Precipitation Data Authors", *Journal of Climate*, 10 (6), pp. 1405-1427.

Fraley, C., and A. E. Raftery (1998). "How many clusters? Which clustering method? Answers via model-based cluster analysis", *The Computer Journal*, 41(8), pp. 578-588.

Fraley, C., and A. E. Raftery (2002a). "Model-based clustering, discriminant analysis and density estimation", *Journal of the American Statistical Association*, 97(458), pp. 611-631.

Fraley, C., and A. E. Raftery (2002b). "MCLUST: Software for model-based clustering, density estimation and discriminant analysis", *Technical Report No. 415*, Department of Statistics, University of Washington.

Freeman, A. M., (1993). *The measurement of environmental and resource values: Theory and methods*, Resources for the Future, Washington, D.C.

Fuller, W. A., (1969). "Grafted polynomials as approximating functions", *Australian Journal of Agricultural Economics*, pp. 35-64.

Gamble, H. B., O. H. Sauerlender and C. J. Langley (1974). "Adverse and beneficial effects of highways on property values", *Transport Research Record*, 508, pp 37-48.

Garrod, G., and K. G. Willis (1992). "Valuing goods' characteristics: an application of the hedonic price method to environmental attributes," *Journal of Environmental management*, 34(1), 59-76.

Gautrin, J.-F., (1975). "An evaluation of the impact of aircraft noise on property values with a simple model of urban land rent", *Land Economics*, 51, pp 80-86.

Gawande, K., and H. Jenkins-Smith (2001). "Nuclear waste transport and residential property values: An hedonic pricing model", *Journal of Environmental Economics and Management*, 42(2), 207-233.

Gençay, R., and X. Yang (1996). "A forecast comparison of residential housing prices by parametric versus semiparametric conditional mean estimators", *Economics Letters*, 52, pp. 129-135.

Gent, van H. A. and P. Rietveld (1993). "Road transport and the environment in Europe", *The Science of the Total Environment*, 129, pp 205-18.

Ghosh, D., and A. M. Chinnaiyan (2002). "Mixture modelling of gene expression data from microarray experiments", *Bioinformatics*, 18, pp. 275-286.

Gibbons, S., (2002). "Paying for good neighbours? Neighbourhood deprivation and the community benefits of education", *Centre for the Economics of Learning Discussion Paper*, 17.

Gibbons, S., (2003). "Paying for good neighbours? Estimating the value of an educated community", *Urban Studies*, 40(1).

Gibbons, S., and S. Machin (2001). "Valuing primary schools", *Centre for the Economics of Learning Discussion Paper*, 15.

Gibbons, S., and S. Machin (2003). "Valuing English primary schools", *Journal of Urban Economics*, 53(2), 15.

Goetzmann, W. N., and S. M. Wachter (1995). "Clustering methods for real estate portfolios", *Real Estate Economics*, 23 (3), pp. 271-310.

Goodman, A. C., (1978). "Hedonic prices, price indexes and housing markets", *Journal of Urban Economics*, 5, pp 471-484.

Goodman, A. C., and R. A. Dubin (1990). "Sample Stratification with Non-Nested Alternatives: Theory and a Hedonic Example," *Review of Economics and Statistics*, 72, pp 168-173.

Goodman, A. C., and T. G. Thibodeau (1998). "Housing market segmentation", *Journal of Housing Economics*, 7, pp. 175-185.

Goodman, A. C., and T. G. Thibodeau (2003). "Housing market segmentation and hedonic prediction accuracy", *Journal of Housing Economics*, 12(3), pp.181-201.

Gopal, S. S., and J. Hebet (1998). "Bayesian Pixel Classification using Spatially Variant Finite Mixures and Generalised EM Algorithm", *IEEE Trans. on Image Processing*, 7, pp. 1014-1028.

Greene, W. H., (1993). *Econometric Analysis*. Second Edition. Macmillan Publishing: New York.

Grue, B., J. L. Langeland and O. I. Larsen (1997). *Housing Prices – impacts of exposure to road traffic and location*, TØI report 351/1997, Oslo.

Hall, F. L., B. E. Breston and S. M. Taylor, (1978). "Effects of highway noise on residential property values", *Transportation Research Record*, 686, pp 38-43.

Hall, F. L., B. E. Breston and S. M. Taylor, (1982). "The effect of road traffic noise on house prices", *Journal of Sound and Vibration*, 80(4), pp 523-541.

Halvorsen, R. and H. O. Pollakowski (1981). "Choice of functional form for hedonic price equations", *Journal of Urban Economics*, 10, pp 37-49.

Hanemann, W. M. (1991). "Willingness to pay and willingness to accept: How much can they differ?", *American Economic Review,* 81(3), pp 635-47.

Härdle, W., and J. S. Marron (1985). "Optimal bandwidth selection in nonparametric regression function estimation", *Annals of Statistics*, 13, pp. 1465-1481.

Harrison, D. and D. L. Rubinfeld (1978). "Hedonic housing prices and the demand for clean air", *Journal of Environmental Economics and Management*, 5, pp 81-102.

Hartigan, J. A. (1975). *Clustering algorithms*. New York: Wiley & Sons.

Hartigan, J. A., and M. A. Wong (1978). "Algorithm AS 136: A *K*-means clustering algorithm", *Applied Statistics*, 28, pp 100-108.

Hastie, T. J., and C. Loader (1993). "Local regression: Automatic kernel carpentry", *Statistical Science*, 8, pp 120-143.

Hausman, J., (1978). "Specification tests in econometrics". *Econometrica*, 46(3), pp. 1251-1271.

Heckman, J., R. Matzkin and L. Nesheim (2002). "Non-parametric estimation of nonadditive hedonic models", unpublished working paper.

Heckman, J., R. Matzkin and L. Nesheim (2003). "Simulation and estimation of hedonic models", *CENMAP Working Paper*, CWP10/03, Institute of Fiscal Studies, UCL, London.

Hepple, L. W., (1998). "Exact testing for spatial autocorrelation among regression residuals", *Environment and Planning, A*, 30, pp 85-108.

Hidano, N., Y. Hayashiyama and M. Inoue, (1992). "Measuring the external effects of noise and vibration of urban transportation by the hedonic approach", *Environmental Science*, 9(3), pp 401-409, Tokyo.

Hoesli, M., and B. D. MacGregor (1995). "The classification of local property markets in the UK using cluster analysis", in *The Cutting Edge: Proceedings of the RICS Property Research Conference, 1995, Vol. 1*, London: Royal Institution of Chartered Surveyors.

Holland, M. R. and N. Eyre (1992). "Evaluation of the external costs of a UK coal fired power station on agricultural crops", *Proceedings of the 2nd International Conference on the External Costs of Electrical Power*, September 1992, Racine, Wisconsin.

Howes, D., and A. Gatrell (1993). "Visibility analysis in GIS: Issues in the environmental impact assessment of windfarm developments", *Proceeding* *of the Fourth European Conference and Exhibition on Geographical Information Systems*. Genoa, Italy.

Huh, S., and S.-J. Kwak (1997). "The choice of functional forma and variables in the hedonic price model in Seoul", *Urban Studies*, 34(7), pp 989-998.

Ihlanfeldt, K. R. and L. O. Taylor (2001). "Externality effects of small-scale hazardous waste sites: evidence from urban commercial property markets" *Environmental Policy Working Paper Series*, #2001-002, Georgia State University, Atlanta.

Institution of Highways and Transportation with the Department of Transport (1987). *Roads and Traffic in Urban Areas*. London: H.M.S.O.

Irwin, E. G., and N. E. Bockstael (1998). "Interacting agents, spatial externalites and the endogenous evolution of land use pattern", *Working Paper 98-24*, Department of Agricultural and Resource Economics, University of Maryland, College Park.

Jain, A. K., and R. C. Dubes (1988). *Algorithms for Clustering Data*. Prentice Hall, Englewood Cliffs, NJ.

Jain, A. K., M. N. Murty and P. J. Flynn (1999). "Data Clustering: A Review", *ACM Computing Surveys*, 31(3), pp. 264-323.

JMP Consultants (1996). *Monetary valuation of road traffic nuisance*, report to the Department of Transport, Job No. 8329.

Johnson, C. R. (1970). "Positive Definite Matrices". American Mathematics Monthly 77, pp. 259-264.

Kaiser, H. F., (1960). "The application of electronic computers to factor analysis", *Educational and Psychological Measurement*, **20**, 141-151.

Kelejian, H. H., and I. R. Prucha (1999). "A generalised moments estimator for the autoregressive parameter in a spatial model", *International Economic Review*, 40(2), pp 509-533.

Kelejian, H. H., and Robinson, D. P., (1992). "Spatial autocorrelation: A new computationally simple test with an application to per capita county police expenditures", *Regional Science and Urban Economics*, 22, pp 317-331.

Keribin, C., (1998). "Consistent estimate of the order of mixture models", *Comptes Rendues de l'Academie des Sciences, série I – Mathématiques*, 326, pp. 243-248. *Annals of Statistics*, 20, pp. 1350-1360.

Kobayashi, M., (1986). "A bounds test for equality between sets of coefficients in two linear regressions when disturbance variances are unequal", *Journal of the American Statistical Association*, 81(394), pp. 510-513.

Krutilla, J. V. and A. C. Fisher (1975). *The Economics of Natural Environments: Studies in the Valuation of Commodity and Amenity Resources*, Johns Hopkins University Press (for Resources for the Future), Baltimore, MD.

Lake, I. R., A. A. Lovett, I. J. Bateman and B. H. Day (2000). "Using GIS and large-scale digital data to implement hedonic pricing studies", *International Journal of Geographical Information Science*, 14(6), pp. 521-541

Lake, I. R., A. A. Lovett, I. J. Bateman and I. H. Langford (1998). "Modelling environmental influences on property prices in an urban environment", *Computers Environment and Urban Systems*, 22, pp. 121-136.

Langley, C. J., (1976). "Adverse impacts of the Washington beltway on residential property values", *Land Economics*, 52, pp 54-65.

Leggett, C. G., and N. E. Bockstael (2000). "Evidence of the effects of water quality on residential land prices", *Journal of Environmental Economics and Management*, 39(2), pp. 121-144.

Leroux, M., (1992). "Consistent estimation of a mixing distribution", *The Annals of Statistics*, 20, pp. 1350-1360.

Levesque, T. J., (1994). "Modelling the effects of airport noise on residential housing markets: A case study of Winnipeg International Airport", *Journal of Transport Economics and Policy*, pp 199-210.

Li, Q., and S. Wang (1998). "A simple consistent bootstrap test for a parametric regression function", Journal of Econometrics, 87, pp 145-165.

Lindeman, R. H., P. F. Merenda and R. Gold (1980). "Chapter 8: Factor Analysis", in *Introduction to Bivariate and Multivariate Analysis*, Scott, Foresman and co.

Litman, T. (1995). *Transportation cost analysis, estimates and implications*, Victoria Transport Policy, B.C.

London Research Centre (1993). *London energy study; Energy use and the environment*, London Research Centre, London.

MacQueen, J., (1967). "Some methods of classification and analysis of multivariate observations." In L. M. L. Cam and J. Neyman (eds.), *Proc. 5th Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, pp. 281-297. University of California Press.

Maddala, G., (1983). *Limited-Dependent and Qualitative Variables in Econometrics,* Cambridge University Press.

Maddison, D., D. Pearce, O. Johansson, E. Calthrop, T. Litman and E. Verhoef (1996). *Blueprint 5. The True Cost of Road Transport*, Earthscan, London.

Maguire, D. J., M. F. Goodchild and D. W. Rhind, (Eds.) (1991). *Geographical information systems; vol.1: principles*. Harlow: Longman Scientific & Technical.

Maitra, R., (2001). "Clustering massive datasets with applications in software metrics and tomography", Technometrics, 20 (3), pp. 336-346.

Martin, D., P. Longley and D. Higgs (1994). "The use of GIS in the analysis of diverse urban datasets." *Computers, Environment and Urban Systems*, 18(1), 55-56.

Maser, S. M., W. H. Riker and R. N. Rosett (1977). "The effects of zoning and externalities on the price of land: An empirical analysis of Monroe County, New York", *Journal of Law and Economics*, 20, pp 111-132.

McConnell, K. E., and T. T. Phipps (1987). "Identification of preference parameters in hedonic models: Consumer demands with nonlinear budgets", *Journal of Urban Economics*, 22, pp 35-52.

McLachlan, G. J., R. W. Bean and D. Peel (2002). "A mixture model-based approach to the clustering of microarray expression data", *Bioinformatics*, 18(3), pp.413-422.

McLachlan, G., (1987). "On bootstrapping the likelihood ratio test statistic for the number of components in a normal mixture", *Applied Statistics*, 36, pp. 318-324.

McLachlan, G., and K. Basford (1988). *Mixture Models*, New York: Marcel Dekker.

McLeod, P.B., (1984). "The demand for local amenity: An hedonic price analysis", *Environment and Planning A*, 16(3), pp 389-400.

McMillan, M. L., B. G. Reid and D. W. Gillen (1978). *An approach towards improved estimates of willingness to pay for public goods from hedonic price functions: A case study of aircraft noise*, unpublished paper, University of Alberta.

Medley, G. (1992). "Nature, the environment and the future", paper presented at *The University of the Third Age, 1992 International Symposium: The Challenges of the Future*, 13th-18th September 1992, Kings College, Cambridge.

Michaels, R. G., and V. K. Smith (1990). "Market-segmentation and valuing amenities with hedonic models - the case of hazardous-waste sites", *Journal of Urban Economics*, 28 (2), pp 223-242.

Mieskowski, P. and A. M. Saper (1978). "An estimate of the effects of airport noise on property values", *Journal of Urban Economics*, 5, pp 425-440.

Millon, J. W., J. Gressel and D. Mulkey (1984). "Hedonic amenity valuation and functional form specification", *Land Economics*, 60 (4), pp 378-387.

Moran, P. A. P., (1950). "Notes on continuous stochastic phenomena", *Biometrika*, 37, pp.17-23.

Mukherjee, S., E. D. Feigelson, G. J. Babu, F. Murtagh, C. Fraley and A. E. Raftery (1998). "Three types of gamma-ray bursts", *Astrophysical Journal*, 508 (1), pp. 314-327.

Murray, M. P., (1982). "Mythical demands and supplies for proper estimation of Rosen's hedonic price model", *Journal of Urban Economics*, 14, pp 327-337.

National Roads Directorate (1998). *The Effect of Road Traffic on Residential Property Values*, Annex I of Project Brief, The Scottish Office Development Department, Edinburgh.

Nelson, J. P., (1978). "Residential choice, hedonic prices and the demand for urban air quality", *Journal of Urban Economics*, 5, pp 357-369.

Nelson, J. P., (1978a). *Economic analysis of transportation noise abatement*, Ballinger, Cambridge, Mass.

Nelson, J. P., (1978b). "Residential choice, hedonic prices and the demand for urban air quality", *Journal of Urban Economics*, 5, pp 357-69

Nelson, J. P., (1979). "Airport Noise, location rent and the market for residential amenities", *Journal of Environmental Economics and Management*, 6, pp 320-331.

Nelson, J. P., (1980). "Airports and property values: A survey of recent evidence", *Journal of Transport Economics and Policy*, 14(1), pp 37-52.

Nelson, J. P., (1982). "Highway noise and property values: A survey of recent evidence", *Journal of Transport Economics and Policy*, 16(2), 117-38.

Nesheim, L., (2002). "Equilibrium sorting of heterogeneous consumers across locations: theory and empirical implications", *CEMMAP Working Paper*, CWP08/02, Insititute of Fiscal Studies, Department of Economics, University College London.

Newey, W., (1987). "Simultaneous estimation of limited dependent variable models with endogenous explanatory variables." *Journal of Econometrics*, 36, pp. 231-250.

O'Byrne, P. H., J. P. Nelson and J. J. Seneca (1985). "Housing values, census estimates, disequilibrium and the environmental costs of airport noise: A case study of Atlanta", *Journal of Environmental Economics and Management*, 12, pp 169-178.

Openshaw, S., (Ed.) (1995). *Census users' handbook*. Cambridge: Pearson Professional Limited.

Opsomer, J., Y. Wang and Y. Yang (2001). "Nonparametric regression with correlated errors", *Statistical Science*, 16(2), pp. 134-153.

Ordnance Survey (1996). *Digital Map Data and Customised Services*. Southampton: Ordnance Survey.

Ostro, B., (1994). *Estimating health effects of air pollution: A methodology with an application to Jakarta*, World Bank, Washington, D.C.

Pace, R. K., (1993). "Nonparametric methods with applications to hedonic models", *Journal of Real Estate Finance and Economics*, 7(3), pp.185-204.

Pace, R. K., (1995). "Parametric, semiparametric and nonparametric estimation of characteristic values with mass assessment and hedonic pricing models", *Journal of Real Estate Finance and Economics*, 11, pp.195-217.

Pace, R. K., and O.W. Gilley (1997). "Using the spatial configuration of the data to improve estimation", *Journal of Real Estate Finance and Economics*, 14, pp.333-340.

Paik, I. K., (1972). *Measurement of environmental externality in particular reference to noise*, unpublished PhD dissertation, Georgetown University.

Palm, R., (1978). "Spatial segmentation of the urban housing market", *Economic Geography*, 54, pp. 210-221.

Palmquist, R. B, and A. Isangkura (1999). "Valuing air quality with hedonic and discrete choice models", *American Journal of Agricultural Economics*, 81(5), pp1128-33.

Palmquist, R. B., (1980). *Impact of highway improvements on property values in Washington*, WA-RD-37.1. Springfield, Va.: National Technical Information Service.

Palmquist, R. B., (1982a). "Measuring environmental effects on property prices values with hedonic regressions", *Journal of Urban Economics*, 11, pp 333-47.

Palmquist, R. B., (1982b). "Estimating the demand for air quality from property value studies", manuscript, Raleigh: North Carolina State University.

Palmquist, R. B., (1984). "Estimating the demand for the characteristics of housing", *Review of Economics and Statistics*, 66, pp 394-404.

Palmquist, R. B., (1988). "Welfare measurement for environmental improvements using the hedonic model: The case of nonparametric marginal prices", *Journal of Environmental Economics and Management*, 15, pp 297-312.

Palmquist, R. B., (1991). "Hedonic methods" in *Measuring the Demand for Environmental Quality*, J. B. Braden and C. D. Kolstad (eds.), Amsterdam: North Holland.

Palmquist, R. B., (1992). "A note on transactions costs, moving costs and benefit measurement." *Journal of Urban Economics*, 32, pp 40-44.

Pan, W., J. Lin and C. T. Le (2002). "Model-based cluster analysis of microarray gene-expression data", *Genome Biology*, 3(2): research 0009.1–0009.8.

Pavlov, A. D., (2000). "Space-varying regression coefficients: A semi-parametric approach applied to real estate markets", *Real Estate Economics*, 28(2), pp 249-283.

Pearce, D. W., and A. Markandya (1989). *The Benefits of Environmental Policy*, Organisation of Economic Cooperation and Development, Paris.

Pearce, D. W., and R. K. Turner (1990) *The Economics of Natural Resources and the Environment*, Harvester Wheatsheaf, Hemel Hempstead.

Pennington, G., N. Topham and R. Ward (1990). "Aircraft noise and residential property values adjacent to Manchester International Airport", *Journal of Transport Economics and Policy*, 24(1), pp 49-59.

Penrose, M. D., (1998). "Extremes for the minimal spanning tree on normally distributed points", *Advances in Applied Probability*, 30, pp 628-639.

Pommerehne, W. W., (1988). "Measuring environmental benefits: Comparison of a hedonic technique and CVM", in *Welfare and Efficiency in Public Economics*, D. Bos, M. Rose and C. Seidl (ed.), Berlin.

Posse, C., (2001). "Hierarchical model-based clustering for large datasets", *Journal Of Computational And Graphical Statistics,* 10 (3), pp. 464-486.

Powell, J. L., (1986). "Symmetrically trimmed least squares estimation for tobit models", *Econometrica*, 54, pp 1435-1460.

Price, I., (1974). *The social cost of airport noise as measured by rental changes: The case of Logan Airport*, unpublished PhD dissertation, Boston University.

Prim, R. C., (1957). "Shortest connection networks and some generalizations", *Bell Systems Technology Journal,* 36, pp. 1389-1401.

Ramussen, D. W., and T. W. Zuehlke (1990). "On the choice of functional form for hedonic price functions", *Applied Economics*, 22, pp 431-438.

Randall, A., and J. R. Stoll (1980). "Consumer's surplus in commodity space", *American Economic Review*, 70(3) pp 449-55.

Renew, W. D., (1996a). *The relationship between traffic noise and house price*, Conference of the Australian Acoustical Society, Brisbane, 13-15th November, 1996.

Renew, W. D., (1996b). *The relationship between traffic house prices and traffic noise in Brisbane*, Draft report, Department of the Environment, Australia.

Ridker, R. A., and J. A. Henning (1967). "The determinants of residential property values with special reference to air pollution", *Rev. Econ. St.*, 49, pp 246-57.

Ridker, R., and J. A. Henning (1976). "The determinants of residential property prices with special reference to air pollution", *Review of Economics and Statistics*, 49, pp 246-257.

Robert, P., and Y. Escoufier (1976). "A unifying tool for linear multivariate statistical methods: the RV-coefficient", *Applied Statistics*, 25, pp. 257-65.

Robinson, P. M., (1988). "Root-*N*-consistent semiparametric regression", *Econometrica*, 56, pp 931-954.

Roeder, K., and L. Wasserman (1997). "Practical Bayesian density estimation using mixtures of normals", *Journal of the American Statistical Association*, 87, pp. 108-119.

Rogers, W. H., (1993). "Regression standard errors in clustered samples", *Stata Technical Bulletin*, 13, pp 19-23.

Rosen, S. (1974). "Hedonic prices and implicit markets: Production differentiation in pure competition", *Journal of Political Economy*, 82, pp 34-55.

Schipper, Y. J. J., (1996). *On the valuation of aircraft noise: A meta-analysis*, unpublished paper, Titenberg Institute, Free University of Amsterdam.

Schnare, A., and R. Struyk (1976). "Segmentation in urban housing markets", *Journal of Urban Economics*, 3, pp 146-166.

Schultz, W., (1986). "A survey on the status of research concerning the evaluation of benefits of environmental policy in the Federal Republic of Germany", paper presented to the *OECD Workshop on the Benefits of Environmental Policy and Decision Making*, Avignon.

Schwartz, G., (1978). "Estimating the dimensions of a model", *Annals of Statistics*, 6, pp. 461-464.

Smith, P. L., (1974). "Splines as a useful and convenient statistical tool", *The American Statistician*, 33, pp. 57-62.

Smith, V. K., and T. A. Deyak (1975). "Measuring the impact of air pollution on property values", *Journal of Regional Science*, 15(3), pp 277-88.

Smyth, P., (2000). "Model selection for probabilistic clustering using cross-validated likelihood", *Statistics and Computing*, 10(1), pp. 63-72.

Smyth, P., K. Ide and M. Ghil (1999). "Multiple regimes in Northern Hemisphere height fields via mixture model clustering", *Journal of Atmospheric Sciences*, 56(21) pp. 3704-3723.

Soguel, N., (1991). *Evaluation de cout social du bruit genere par le trafic routier en Ville de Neuchatel*, IRER, WP N°. 9105, Universite de Neuchatel.

Soguel, N., (1994). *Evaluation monetaire des atteinties a l'environment: Une etude hedoniste et contingente sur l'impact des transports*, Imprimerie de L'evolve SA Neuchatel.

Sonstelie, J. C., and P. R. Portney (1980). "Gross rents and market values: Testing the implications of Trebout's hypothesis", *Journal of Urban Economics*, 7, pp 102-118.

Stanford, D., and A. E. Raftery (2000). "Principal curve clustering with noise", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22, pp. 601-609.

Straszheim, M. R., (1973). "Estimation of the demand for urban housing services from household interview data", *Review of Economics and Statistics*, 55, pp 1-8.

Straszheim, M. R., (1974). "Hedonic estimation of housing market prices: A further comment", *Review of Economics and Statistics*, 56, pp 404-06.

Straszheim, M. R., (1975). *An Econometric Analysis of the Urban Housing Market*, New York: National Bureau of Economic Research.

Sugden, R., (1999). "Alternatives to the Neoclassical Theory of Choice", Chapter 6 in *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries* (I. Bateman and K.G. Willis, eds.), Oxford University Press.

Tauchen, H., and A. D. Witte (2001). "Estimating hedonic models: Implications of the theory", National Bureau of Economic Research, Technical Working Paper, 271.

Ter Braak, C. J. F., H. Hoijtink, W. Akkermans, P. F. M. Verdonschot (2003). "Bayesian model-based cluster analysis for predicting macrofaunal communities", *Ecological Modelling*, 160 (3), pp. 235-248.

Tiao, G.C., and A. Goldberger (1962). "Testing equality of individual regression co-efficients", *WEBH Paper 6201*, University of Wisconsin, Social Systems Research Institute.

Tinbergen, J., (1956). "On the theory of income distribution", Weltwirtschaftliches Archiv, 77, pp 155-73.

Tinch, R., (1996). "Valuation of Environmental Externalities", Transport Policy Unit, Department of Transport (former), London.

Titterington, D. M., A. F. M. Smith and U. E. Markov (1985). *Statistical Analysis of Finite Mixture Distributions*, John Wiley and Sons, Chichester, UK.

Toyoda, T., (1974). "Use of the Chow test under heteroskedasticity", *Econometrica*, 42, pp. 601–608.

Uyeno, D., S. W. Hamilton and A. Biggs (1993). "Density of residential land use and the impact of airport noise", *Journal of Transport and Economics and Policy*, 27(1), pp 3-18.

Vainio, M., (1995). *Traffic noise and air pollution: Valuation of Externalities with the hedonic price and contingent valuation methods*, PhD Thesis, School of Economics and Business Administration, Helsinki.

Vaughan, R. J., and L. Huckins (1975). *The economics of expressway noise and pollution abatement*, P-5475, Santa Monica, Ca.: Ballinger.

Walters, A. A., (1975). *Noise and prices*, Oxford University Press, London.

Wand, M. P., (1994). "Fast implementations of nonparametric curve estimators", *Journal of Computational and Graphical Statistics*, 3(4), pp. 433-445.

Ward, J. H., (1963). "Hierarchical groupings to optimize an objective function", *Journal of the American Statistical Association*, 58, pp. 234–244.

Watkins, C. A., (2001). "The definition and identification of housing submarkets", *Environment and Planning A*, 33, pp. 2235-2253.

Webster, R., and M. A. Oliver (1990) *Statistical Methods in Soil and Land Resource Survey*, Oxford University Press, Oxford.

Wehrens, R., L. M. C. Buydens, C. Fraley and A. E. Raftery (2003). "Model-based clustering for image segmentation and large datasets via sampling", *Technical Report No. 424*, Department of Statistics, University of Washington.

Weisbrod, B. A., (1964). "Collective - consumption services of individual consumption goods", *Quarterly Journal of Economics*, 78:471-477.

Whang, Y-J, and D. W. K. Andrews (1993). "Tests of specification for parametric and semiparametric models", *Journal of Econometrics*, 57, pp 277-318.

Wieand, K. F., (1973). "Air pollution and property values: A study of the St. Louis area", *Journal of Regional Science*, 13, pp 91-95.

Williams, R. L., (2000). "A note on robust variance estimation for cluster-correlated data", *Biometrics*, 56, pp 645-646.

Willig, R. D., (1973). "Consumers surplus: a rigorous cookbook", *Technical Report No. 98*, Institute for Mathematical Studies in the Social Sciences, Stanford University.

Willig, R. D., (1976). "Consumer's surplus without apology", *American Economic Review*, 66(4), pp 589-597.

Wolverton, M. L., W. G. Hardin and P. Cheng (1999). "Disaggregation of local apartment markets by unit type", *Journal of Real Estate Finance and Economics*, 19 (3), pp. 243-257.

Xu, D., V. Olman , L. Wang and Y. Xu (2003). "EXCAVATOR: a computer program for efficiently mining gene expression data", Nucleid Acids Research, 31 (19), pp. 5582-5589.

Xu, Y., V. Olman and D. Xu (2002). "Clustering gene expression data using a graph-theoretic approach: an application of minimum spanning trees", *Bioinformatics*, 18 (4), pp. 536-545.

Yamaguchi, Y., (1996). "Estimating the cost of aircraft noise around airports in London", *unpublished MSc dissertation*, University College London.

Yeung, K. Y., C. Fraley, A. Murua, A. E. Raftery, W. L. Ruzzo (2001). "Model-based clustering and data transformations for gene expression data", *Bioinformatics*, 17, pp. 977-987.

Zabel, J. E. and K. A. Kiel (2000). "Estimating the demand for air quality in four U.S. cities", *Land Economics*, 76(2), pp 174-194.